

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
5 June 2003 (05.06.2003)

PCT

(10) International Publication Number  
WO 03/046152 A2

(51) International Patent Classification<sup>7</sup>: C12N

(21) International Application Number: PCT/US02/38446

(22) International Filing Date:  
25 November 2002 (25.11.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/334,182	28 November 2001 (28.11.2001)	US
60/342,052	18 December 2001 (18.12.2001)	US
60/350,410	18 January 2002 (18.01.2002)	US
60/353,284	1 February 2002 (01.02.2002)	US
60/363,649	8 March 2002 (08.03.2002)	US

(71) Applicant (for all designated States except US): **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **CHAWLA, Narinder, K.** [US/US]; 33 Union Square, #712, Union City, CA 94587 (US). **LU, Dyung Aina, M.** [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). **KHAN, Farrah, A.** [IN/US]; 3617 Central Road #102, Glenview, IL 60025 (US). **GANDHI, Ameena, R.** [US/US]; 705 5th Avenue, San Francisco, CA 94118 (US). **SWARNAKAR, Anita** [CA/US]; 8 Locksley Avenue #5D, San Francisco, CA 94122 (US). **AZIMZAI, Yalda** [US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA 94552 (US). **MARQUIS, Joseph, P.** [US/US]; 4428 Lazy Lane, San Jose, CA 95135 (US). **SPRAGUE, William, W.** [US/US]; 611 13th Street # C, Sacramento, CA 95814 (US). **EMERLING, Brooke, M.** [US/US]; 1735 Woodland Avenue # 71, Palo Alto, CA 94303 (US). **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **BOROWSKY, Mark, L.** [US/US]; 122 Orchard Avenue, Redwood City, CA 94061 (US). **BECHA, Shanya, D.** [US/US]; 21062 Gary Drive # 117, Castro Valley, CA 94546 (US). **ISON, Craig, H.** [US/US]; 1242 Weathersfield Way, San Jose, CA 95118 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US). **HAFALIA, April, J., A.** [US/US]; 2227

Calle de Primavera, Santa Clara, CA 95054 (US). **RING, Huijun, Z.** [US/US]; 625 Orange Avenue, Los Altos, CA 94022 (US). **WARREN, Bridget, A.** [US/US]; 1810 S. El Camino Real #B103, Encinitas, CA 94024 (US). **GIETZEN, Kimberly, J.** [US/US]; 691 Los Huecos Drive, San Jose, CA 95123 (US). **TRAN, Uyen, K.** [US/US]; 2638 Mabury Square, San Jose, CA 95133 (US). **LEE, Soo, Yeun** [KR/US]; 40 Westdale Avenue, Daly City, CA 94015 (US). **LEE, Ernestine, A.** [US/US]; 20523 Crow Creek Road, Castro Valley, CA 94552 (US). **RICHARDSON, Thomas, W.** [US/US]; 616 Canyon Road #107, Redwood City, CA 94062 (US). **KABLE, Amy, E.** [US/US]; 2345 Polk Street #4, San Francisco, CA 94109 (US). **BURFORD, Neil** [GB/US]; 105 Wildwood Circle, Durham, CT 06422 (US). **LEHR-MASON, Patricia, M.** [US/US]; 360 Clarke Lane, Morgan Hill, CA 95014 (US). **GORVAD, Ann, E.** [US/US]; 369 Marie Common, Livermore, CA 94550 (US). **LEE, Sally** [US/US]; 3643 26th Street, San Francisco, CA 94110 (US). **BLAKE, Julie, J.** [US/US]; 3818 Pacheco Street, San Francisco, CA 94116 (US). **HONCHELL, Cynthia, D.** [US/US]; 158 Laurel Street, San Carlos, CA 94070 (US). **THANGAVELU, Kavitha** [IN/US]; 1950 Montecito Avenue #23, Mountain View, CA 94043 (US). **RAMKUMAR, Jayalaxmi** [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). **CHIEN, David** [US/US]; 444 Yerba Buena Avenue, Los Altos, CA 94022 (US). **JIN, Pei** [US/US]; 320 Curtner Avenue #D, Palo Alto, CA 94306 (US). **CHANG, Hsin-Ru** [US/US]; 326 Treasure Island Drive, Belmont, CA 94002 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **NGUYEN, Danniell, B.** [US/US]; 4922 Arnica Court, San Jose, CA 95111 (US). **KHARE, Reena** [US/US]; 12650 Orella Court, Saratoga, CA 95070 (US). **BHATIA, Umesh** [US/US]; 5212 Union Avenue, San Jose, CA 95124 (US). **BURRILL, John, D.** [US/US]; 2218 Brewster Avenue, Redwood City, CA 94062 (US). **HO, Anne** [FR/US]; 1279 Poplar Avenue, #114, Sunnyvale, CA 94086 (US). **ZHENG, Wenjin** [CN/US]; 9 Sutter Creek Lane, Mountain View, CA 94043 (US).

(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,

[Continued on next page]

(54) Title: MOLECULES FOR DISEASE DETECTION AND TREATMENT

(57) Abstract: Various embodiments of the invention provide human molecules for diseasedetection and treatment (MDDT) and polynucleotides which identify and encode MDDT. Embodiments of the invention also provide expression vectors, host cells, antibodies, agonists, and antagonists. Other embodiments provide methods for diagnosing, treating, or preventing disorders associated with aberrant expression of MDDT.



LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PI, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— *without international search report and to be republished upon receipt of that report*

**(84) Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK,

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

## MOLECULES FOR DISEASE DETECTION AND TREATMENT

### TECHNICAL FIELD

The invention relates to novel nucleic acids, molecules for disease detection and treatment encoded by these nucleic acids, and to the use of these nucleic acids and proteins in the diagnosis, treatment, and prevention of cell proliferative, autoimmune/inflammatory, developmental, and neurological disorders. The invention also relates to the assessment of the effects of exogenous compounds on the expression of nucleic acids and molecules for disease detection and treatment.

### BACKGROUND OF THE INVENTION

It is estimated that only 2% of mammalian DNA encodes proteins, and only a small fraction of the genes that encode proteins are actually expressed in a particular cell at any time. The various types of cells in a multicellular organism differ dramatically both in structure and function, and the identity of a particular cell is conferred by its unique pattern of gene expression. In addition, different cell types express overlapping but distinctive sets of genes throughout development. Cell growth and proliferation, cell differentiation, the immune response, apoptosis, and other processes that contribute to organismal development and survival are governed by regulation of gene expression. Appropriate gene regulation also ensures that cells function efficiently by expressing only those genes whose functions are required at a given time. Factors that influence gene expression include extracellular signals that mediate cell-cell communication and coordinate the activities of different cell types. Gene expression is regulated at the level of DNA and RNA transcription, and at the level of mRNA translation.

Aberrant expression or mutations in genes and their products may cause, or increase susceptibility to, a variety of human diseases such as cancer and other cell proliferative disorders. The identification of these genes and their products is the basis of an ever-expanding effort to find markers for early detection of diseases and targets for their prevention and treatment. For example, cancer represents a type of cell proliferative disorder that affects nearly every tissue in the body. The development of cancer, or oncogenesis, is often correlated with the conversion of a normal gene into a cancer-causing gene, or oncogene, through abnormal expression or mutation. Oncoproteins, the products of oncogenes, include a variety of molecules that influence cell proliferation, such as growth factors, growth factor receptors, intracellular signal transducers, nuclear transcription factors, and cell-cycle control proteins. In contrast, tumor-suppressor genes are involved in inhibiting cell proliferation. Mutations which reduce or abrogate the function of tumor-suppressor genes result in aberrant cell proliferation and cancer. Thus a wide variety of genes and their products have been found that are associated with cell proliferative disorders such as cancer, but many more may exist

that are yet to be discovered.

DNA-based arrays can provide an efficient, high-throughput method to examine gene expression and genetic variability. For example, SNPs, or single nucleotide polymorphisms, are the most common type of human genetic variation. DNA-based arrays can dramatically accelerate the discovery of SNPs in hundreds and even thousands of genes. Likewise, such arrays can be used for SNP genotyping in which DNA samples from individuals or populations are assayed for the presence of selected SNPs. These approaches will ultimately lead to the systematic identification of all genetic variations in the human genome and the correlation of certain genetic variations with disease susceptibility, responsiveness to drug treatments, and other medically relevant information. (See, for example, Wang, D.G. et al. (1998) *Science* 280:1077-1082.)

DNA-based array technology is especially important for the rapid analysis of global gene expression patterns. For example, genetic predisposition, disease, or therapeutic treatment may directly or indirectly affect the expression of a large number of genes in a given tissue. In this case, it is useful to develop a profile, or transcript image, of all the genes that are expressed and the levels at which they are expressed in that particular tissue. A profile generated from an individual or population affected with a certain disease or undergoing a particular therapy may be compared with a profile generated from a control individual or population. Such analysis does not require knowledge of gene function, as the expression profiles can be subjected to mathematical analyses which simply treat each gene as a marker. Furthermore, gene expression profiles may help dissect biological pathways by identifying all the genes expressed, for example, at a certain developmental stage, in a particular tissue, or in response to disease or treatment. (See, for example, Lander, E.S. et al. (1996) *Science* 274:536-539.)

Certain genes are known to be associated with diseases because of their chromosomal location, such as the genes in the myotonic dystrophy (DM) regions of mouse and human. The mutation underlying DM has been localized to a gene encoding the DM-kinase protein, but another active gene, DMR-N9, is in close proximity to the DM-kinase gene (Jansen, G. et al. (1992) *Nat. Genet.* 1:261-266). DMR-N9 encodes a 650 amino acid protein that contains WD repeats, motifs found in cell signaling proteins. DMR-N9 is expressed in all neural tissues and in the testis, suggesting a role for DMR-N9 in the manifestation of mental and testicular symptoms in severe cases of DM (Jansen, G. et al. (1995) *Hum. Mol. Genet.* 4:843-852).

Other genes are identified based upon their expression patterns or association with disease syndromes. For example, autoantibodies to subcellular organelles are found in patients with systemic rheumatic diseases. A recently identified protein, golgin-67, belongs to a family of Golgi autoantigens having alpha-helical coiled-coil domains (Eystathioy, T. et al. (2000) *J. Autoimmun.* 14:179-187). The Stac gene was identified as a brain specific, developmentally regulated gene. The



Stac protein contains an SH3 domain, and is thought to be involved in neuron-specific signal transduction (Suzuki, H. et al. (1996) Biochem. Biophys. Res. Commun. 229:902-909).

#### Osteoarthritis:

Osteoarthritis (OA) is a debilitating joint disease involving focal cartilage loss. Several studies indicate a major genetic component can be involved in causing OA. Estimates of inheritability from twin studies of radiographic OA of the hand, knee and hip range from 36% to 68% (MacGregor, A.J. and Spector, T.D. (1999) Rheumatology 38:583-560). Several interleukin and interleukin-associated genes are located at 2q12-q22 (Leppavouri, J. et al. (1999) Am. J. Hum. Genet. 65:1060-1067). Interleukins regulate a number of enzymes that degrade the cartilage extracellular matrix, and the expression of certain interleukin genes, including IL-1 $\beta$ , is altered in OA joint tissue (Elson, C.J. et al. (1998) Br. J. Rheum. 37:106-107).

#### Lung Cancer:

Lung cancer is the leading cause of cancer death in the United States, affecting more than 100,000 men and 50,000 women each year. Nearly 90% of the patients diagnosed with lung cancer are cigarette smokers. Tobacco smoke contains thousands of noxious substances that induce carcinogen metabolizing enzymes and covalent DNA adduct formation in the exposed bronchial epithelium. In nearly 80% of patients diagnosed with lung cancer, metastasis has already occurred. Most commonly lung cancers metastasize to pleura, brain, bone, pericardium, and liver. This adversely affects the overall five-year survival rate which is 37% for squamous carcinoma, 27% for adenocarcinoma and large cell carcinoma, and less than 1% for small cell carcinomas. Earlier diagnosis and an systematic approach to identification, staging, and treatment could positively affect patient outcome (DeVita et al. (1997) Cancer: Principles and Practice of Oncology, Lippincott-Raven, Philadelphia PA) and Fauci et al. (1998) Harrison's Principles of Internal Medicine, McGraw Hill, New York, NY).

Lung cancers progress through a series of morphologically distinct stages from hyperplasia to invasive carcinoma. Malignant lung cancers are divided into two groups comprising four histopathological classes. The nonsmall cell lung carcinoma (NSCLC) group includes squamous cell carcinomas, adenocarcinomas, and large cell carcinomas and accounts for about 70% of all lung cancer cases. Adenocarcinomas typically arise in the peripheral airways and often form mucin secreting glands. Squamous cell carcinomas typically arise in proximal airways. The histogenesis of squamous cell carcinomas may be related to chronic inflammation and injury to the bronchial epithelium, leading to squamous metaplasia. The small cell lung carcinoma (SCLC) group accounts for about 20% of lung cancer cases. SCLCs typically arise in proximal airways and exhibit a number of paraneoplastic syndromes including inappropriate

production of adrenocorticotropin and anti-diuretic hormone.

Lung cancer cells accumulate numerous genetic lesions, many of which are associated with cytologically visible chromosomal aberrations. The high frequency of chromosomal deletions associated with lung cancer may reflect the role of multiple tumor suppressor loci in the etiology of this disease. Several studies report deletions of regions of chromosome 11 in NSCLC (Bepler, G. and Garcia-Blanco, M.A. (1994) PNAS 91:5513-7; Iizuka, M., et al. (1995) Genes, Chromosomes & Cancer 13:40-46; Rasio, D. (1995) Cancer Research 55:3988-91). Deletions in other chromosome arms such as 3p, 9p and 17p are also common. Other frequently observed genetic lesions include overexpression of telomerase, activation of oncogenes such as K-ras and c-myc, and inactivation of tumor suppressor genes such as RB, p53 and p16 (Toomey, D. et al. (2001) Cancer 92:2648-57; Zajac-Kaye M. (2001) Lung Cancer 34:S43-6; Wright, G. et al. (2000) Current Opinion in Oncology 12:143-8; Kohno, T. and Yokota, J. (1999) Carcinogenesis 20:1403-10).

#### Parkinson's disease

Parkinson's disease (PD) is a neurodegenerative disorder characterized by the progressive degeneration of the dopaminergic nigrostriatal pathway, and the presence of Lewy bodies. Genetic linkages to chromosomes 2p4, 4p5, and three loci on 1q6-8 have been identified (Gwinn-Hardy K. (2002) Mov. Disord. 17:645-656). Clinical disorders classified as parkinsonism include PD, dementia with Lewy bodies (DLB), progressive supranuclear palsy (PSP), and essential tremor. Several neurodegenerative diseases share pathogenic mechanisms involving tau or synuclein aggregation. These disorders include Alzheimer's disease, and Pick's disease as well as PD and progressive supranuclear palsy (Hardy, J. (2001) J. Alzheimers Dis. 3:109-116). Several genetically distinct forms of PD can be caused by mutations in single genes. Genes for monogenically inherited forms of Parkinson's disease (PD) have been mapped and/or cloned. In some families with autosomal dominant inheritance and typical Lewy-body pathology, mutations have been identified in the gene for alpha-synuclein. Aggregation of this protein in Lewy-bodies may be a crucial step in the molecular pathogenesis of familial and sporadic PD. On the other hand, mutations in the parkin gene cause early-onset autosomal recessive parkinsonism in which nigral degeneration is not accompanied by Lewy-body formation. Parkin-mutations appear to be a common cause of PD in patients with very early onset. Parkin has been implicated in the cellular protein degradation pathways, as it has been shown that it functions as a ubiquitin ligase. A mutation in the gene for ubiquitin C-terminal hydrolase L1 in this pathway has been identified in another small family with PD. Other loci have been mapped to chromosome 2p and 4p, respectively, in families with dominantly inherited PD. These early-onset forms differ from the common sporadic form of PD. It is widely believed that a combination of interacting genetic and environmental causes may be responsible in the majority of

PD-cases (Gasser, T. (2001) J. Neurol. 2001 248:833-840).

#### Expression profiling

Microarrays are analytical tools used in bioanalysis. A microarray has a plurality of molecules spatially distributed over, and stably associated with, the surface of a solid support.

5 Microarrays of polypeptides, polynucleotides, and/or antibodies have been developed and find use in a variety of applications, such as gene sequencing, monitoring gene expression, gene mapping, bacterial identification, drug discovery, and combinatorial chemistry.

One area in particular in which microarrays find use is in gene expression analysis. Array technology can provide a simple way to explore the expression of a single polymorphic gene or the  
10 expression profile of a large number of related or unrelated genes. When the expression of a single gene is examined, arrays are employed to detect the expression of a specific gene or its variants. When an expression profile is examined, arrays provide a platform for identifying genes that are tissue specific, are affected by a substance being tested in a toxicology assay, are part of a signaling cascade, carry out housekeeping functions, or are specifically related to a particular genetic  
15 predisposition, condition, disease, or disorder.

#### Breast Cancer

There are more than 180,000 new cases of breast cancer diagnosed each year, and the mortality rate for breast cancer approaches 10% of all deaths in females between the ages of 45-54 (K. Gish (1999) AWIS Magazine 28:7-10). However the survival rate based on early diagnosis of  
20 localized breast cancer is extremely high (97%), compared with the advanced stage of the disease in which the tumor has spread beyond the breast (22%). Current procedures for clinical breast examination are lacking in sensitivity and specificity, and efforts are underway to develop comprehensive gene expression profiles for breast cancer that may be used in conjunction with conventional screening methods to improve diagnosis and prognosis of this disease (Perou CM et al.  
25 (2000) Nature 406:747-752).

Breast cancer is a genetic disease commonly caused by mutations in cellular disease. Mutations in two genes, BRCA1 and BRCA2, are known to greatly predispose a woman to breast cancer and may be passed on from parents to children (Gish, supra). However, this type of hereditary breast cancer accounts for only about 5% to 9% of breast cancers, while the vast majority of breast  
30 cancer is due to noninherited mutations that occur in breast epithelial cells.

A good deal is already known about the expression of specific genes associated with breast cancer. For example, the relationship between expression of epidermal growth factor (EGF) and its receptor, EGFR, to human mammary carcinoma has been particularly well studied. (See Khazaie et al., supra, and references cited therein for a review of this area.) Overexpression of EGFR,  
35 particularly coupled with down-regulation of the estrogen receptor, is a marker of poor prognosis in

breast cancer patients. In addition, EGFR expression in breast tumor metastases is frequently elevated relative to the primary tumor, suggesting that EGFR is involved in tumor progression and metastasis. This is supported by accumulating evidence that EGF has effects on cell functions related to metastatic potential, such as cell motility, chemotaxis, secretion and differentiation. Changes in expression of other members of the erbB receptor family, of which EGFR is one, have also been implicated in breast cancer. The abundance of erbB receptors, such as HER-2/neu, HER-3, and HER-4, and their ligands in breast cancer points to their functional importance in the pathogenesis of the disease, and may therefore provide targets for therapy of the disease (Bacus, SS et al. (1994) *Am J Clin Pathol* 102:S13-S24). Other known markers of breast cancer include a human secreted frizzled protein mRNA that is downregulated in breast tumors; the matrix Gla protein which is overexpressed in human breast carcinoma cells; Drg1 or RTP, a gene whose expression is diminished in colon, breast, and prostate tumors; maspin, a tumor suppressor gene downregulated in invasive breast carcinomas; and CaN19, a member of the S100 protein family, all of which are down regulated in mammary carcinoma cells relative to normal mammary epithelial cells (Zhou Z et al. (1998) *Int J Cancer* 78:95-99; Chen, L et al. (1990) *Oncogene* 5:1391-1395; Ulrix W et al. (1999) *FEBS Lett* 455:23-26; Sager, R et al. (1996) *Curr Top Microbiol Immunol* 213:51-64; and Lee, SW et al. (1992) *Proc Natl Acad Sci USA* 89:2504-2508).

Cell lines derived from human mammary epithelial cells at various stages of breast cancer provide a useful model to study the process of malignant transformation and tumor progression as it has been shown that these cell lines retain many of the properties of their parental tumors for lengthy culture periods (Wistuba II et al. (1998) *Clin Cancer Res* 4:2931-2938). Such a model is particularly useful for comparing phenotypic and molecular characteristics of human mammary epithelial cells at various stages of malignant transformation.

#### Genes Expressed in C3a Liver Cell Cultures Treated with Steroids

The potential application of gene expression profiling is particularly relevant to measuring the toxic response to potential therapeutic compounds and of the metabolic response to therapeutic agents. Diseases treated with steroids and disorders caused by the metabolic response to treatment with steroids include adenomatosis, cholestasis, cirrhosis, hemangioma, Henoch-Schonlein purpura, hepatitis, hepatocellular and metastatic carcinomas, idiopathic thrombocytopenic purpura, porphyria, sarcoidosis, and Wilson disease. Response may be measured by comparing both the levels and sequences expressed in tissues from subjects exposed to or treated with steroid compounds such as mifepristone, progesterone, beclomethasone, medroxyprogesterone, budesonide, prednisone, dexamethasone, betamethasone, or danazol with the levels and sequences expressed in normal untreated tissue.

Steroids are a class of lipid-soluble molecules, including cholesterol, bile acids, vitamin D,

and hormones, that share a common four-ring structure based on cyclopentanoperhydrophenanthrene and that carry out a wide variety of functions. Cholesterol, for example, is a component of cell membranes that controls membrane fluidity. It is also a precursor for bile acids which solubilize lipids and facilitate absorption in the small intestine during digestion. Vitamin D regulates the absorption of calcium in the small intestine and controls the concentration of calcium in plasma. Steroid hormones, produced by the adrenal cortex, ovaries, and testes, include glucocorticoids, mineralocorticoids, androgens, and estrogens. They control various biological processes by binding to intracellular receptors that regulate transcription of specific genes in the nucleus. Glucocorticoids, for example, increase blood glucose concentrations by regulation of gluconeogenesis in the liver, increase blood concentrations of fatty acids by promoting lipolysis in adipose tissues, modulate sensitivity to catecholamines in the central nervous system, and reduce inflammation. The principal mineralocorticoid, aldosterone, is produced by the adrenal cortex and acts on cells of the distal tubules of the kidney to enhance sodium ion reabsorption. Androgens, produced by the interstitial cells of Leydig in the testis, include the male sex hormone testosterone, which triggers changes at puberty, the production of sperm and maintenance of secondary sexual characteristics. Female sex hormones, estrogen and progesterone, are produced by the ovaries and also by the placenta and adrenal cortex of the fetus during pregnancy. Estrogen regulates female reproductive processes and secondary sexual characteristics. Progesterone regulates changes in the endometrium during the menstrual cycle and pregnancy.

Steroid hormones are widely used for fertility control and in anti-inflammatory treatments for physical injuries and diseases such as arthritis, asthma, and auto-immune disorders. Progesterone, a naturally occurring progestin, is primarily used to treat amenorrhea, abnormal uterine bleeding, or as a contraceptive. Endogenous progesterone is responsible for inducing secretory activity in the endometrium of the estrogen-primed uterus in preparation for the implantation of a fertilized egg and for the maintenance of pregnancy. It is secreted from the corpus luteum in response to luteinizing hormone (LH). The primary contraceptive effect of exogenous progestins involves the suppression of the midcycle surge of LH. At the cellular level, progestins diffuse freely into target cells and bind to the progesterone receptor. Target cells include the female reproductive tract, the mammary gland, the hypothalamus, and the pituitary. Once bound to the receptor, progestins slow the frequency of release of gonadotropin releasing hormone from the hypothalamus and blunt the pre-ovulatory LH surge, thereby preventing follicular maturation and ovulation. Progesterone has minimal estrogenic and androgenic activity. Progesterone is metabolized hepatically to pregnanediol and conjugated with glucuronic acid.

Medroxyprogesterone (MAH), also known as 6 $\alpha$ -methyl-17-hydroxyprogesterone, is a synthetic progestin with a pharmacological activity about 15 times greater than progesterone. MAH

is used for the treatment of renal and endometrial carcinomas, amenorrhea, abnormal uterine bleeding, and endometriosis associated with hormonal imbalance. MAH has a stimulatory effect on respiratory centers and has been used in cases of low blood oxygenation caused by sleep apnea, chronic obstructive pulmonary disease, or hypercapnia.

5 Mifepristone, also known as RU-486, is an antiprogesterone drug that blocks receptors of progesterone. It counteracts the effects of progesterone, which is needed to sustain pregnancy. Mifepristone induces spontaneous abortion when administered in early pregnancy followed by treatment with the prostaglandin, misoprostol. Further, studies show that mifepristone at a substantially lower dose can be highly effective as a postcoital contraceptive when administered  
10 within five days after unprotected intercourse, thus providing women with a "morning-after pill" in case of contraceptive failure or sexual assault. Mifepristone also has potential uses in the treatment of breast and ovarian cancers in cases in which tumors are progesterone-dependent. It interferes with steroid-dependent growth of brain meningiomas, and may be useful in treatment of endometriosis where it blocks the estrogen-dependent growth of endometrial tissues. It may also be useful in  
15 treatment of uterine fibroid tumors and Cushing's Syndrome. Mifepristone binds to glucocorticoid receptors and interferes with cortisol binding. Mifepristone also may act as an anti-glucocorticoid and be effective for treating conditions where cortisol levels are elevated such as AIDS, anorexia nervosa, ulcers, diabetes, Parkinson's disease, multiple sclerosis, and Alzheimer's disease.

Danazol is a synthetic steroid derived from ethinyl testosterone. Danazol indirectly reduces  
20 estrogen production by lowering pituitary synthesis of follicle-stimulating hormone and LH. Danazol also binds to sex hormone receptors in target tissues, thereby exhibiting anabolic, antiestrogenic, and weakly androgenic activity. Danazol does not possess any progestogenic activity, and does not suppress normal pituitary release of corticotropin or release of cortisol by the adrenal glands. Danazol is used in the treatment of endometriosis to relieve pain and inhibit endometrial cell growth.  
25 It is also used to treat fibrocystic breast disease and hereditary angioedema.

Corticosteroids are used to relieve inflammation and to suppress the immune response. They inhibit eosinophil, basophil, and airway epithelial cell function by regulation of cytokines that mediate the inflammatory response. They inhibit leukocyte infiltration at the site of inflammation, interfere in the function of mediators of the inflammatory response, and suppress the humoral  
30 immune response. Corticosteroids are used to treat allergies, asthma, arthritis, and skin conditions. Beclomethasone is a synthetic glucocorticoid that is used to treat steroid-dependent asthma, to relieve symptoms associated with allergic or nonallergic (vasomotor) rhinitis, or to prevent recurrent nasal polyps following surgical removal. The anti-inflammatory and vasoconstrictive effects of intranasal beclomethasone are 5000 times greater than those produced by hydrocortisone. Budesonide is a  
35 corticosteroid used to control symptoms associated with allergic rhinitis or asthma. Budesonide has

high topical anti-inflammatory activity but low systemic activity. Dexamethasone is a synthetic glucocorticoid used in anti-inflammatory or immunosuppressive compositions. It is also used in inhalants to prevent symptoms of asthma. Due to its greater ability to reach the central nervous system, dexamethasone is usually the treatment of choice to control cerebral edema. Dexamethasone is approximately 20-30 times more potent than hydrocortisone and 5-7 times more potent than prednisone. Prednisone is metabolized in the liver to its active form, prednisolone, a glucocorticoid with anti-inflammatory properties. Prednisone is approximately 4 times more potent than hydrocortisone and the duration of action of prednisone is intermediate between hydrocortisone and dexamethasone. Prednisone is used to treat allograft rejection, asthma, systemic lupus erythematosus, arthritis, ulcerative colitis, and other inflammatory conditions. Betamethasone is a synthetic glucocorticoid with antiinflammatory and immunosuppressive activity and is used to treat psoriasis and fungal infections, such as athlete's foot and ringworm.

The anti-inflammatory actions of corticosteroids are thought to involve phospholipase A<sub>2</sub> inhibitory proteins, collectively called lipocortins. Lipocortins, in turn, control the biosynthesis of potent mediators of inflammation such as prostaglandins and leukotrienes by inhibiting the release of the precursor molecule arachidonic acid. Proposed mechanisms of action include decreased IgE synthesis, increased number of  $\beta$ -adrenergic receptors on leukocytes, and decreased arachidonic acid metabolism. During an immediate allergic reaction, such as in chronic bronchial asthma, allergens bridge the IgE antibodies on the surface of mast cells, which triggers these cells to release chemotactic substances. Mast cell influx and activation, therefore, is partially responsible for the inflammation and hyperirritability of the oral mucosa in asthmatic patients. This inflammation can be retarded by administration of corticosteroids.

The effects upon liver metabolism and hormone clearance mechanisms are important to understand the pharmacodynamics of a drug. The human C3A cell line is a clonal derivative of HepG2/C3 (hepatoma cell line, isolated from a 15-year-old male with liver tumor), which was selected for strong contact inhibition of growth. The use of a clonal population enhances the reproducibility of the cells. C3A cells have many characteristics of primary human hepatocytes in culture: i) expression of insulin receptor and insulin-like growth factor II receptor; ii) secretion of a high ratio of serum albumin compared with  $\alpha$ -fetoprotein iii) conversion of ammonia to urea and glutamine; iv) metabolize aromatic amino acids; and v) proliferate in glucose-free and insulin-free medium. The C3A cell line is now well established as an *in vitro* model of the mature human liver (Mickelson et al. (1995) Hepatology 22:866-875; Nagendra et al. (1997) Am J Physiol 272:G408-G416).

#### Colon Cancer:

Colon cancer evolves through a multi-step process whereby pre-malignant colonocytes

undergo a relatively defined sequence of events leading to tumor formation. Several factors participate in the process of tumor progression and malignant transformation including genetic factors, mutations, and selection.

#### Alzheimer's Disease:

Alzheimer's disease is a progressive neurodegenerative disorder that is characterized by the formation of senile plaques and neurofibrillary tangles containing amyloid beta peptide. These plaques are found in limbic and association cortices of the brain. The hippocampus is part of the limbic system and plays an important role in learning and memory. In subjects with Alzheimer's disease, accumulating plaques damage the neuronal architecture in limbic areas and eventually cripple the memory process.

There is a need in the art for new compositions, including nucleic acids and proteins, for the diagnosis, prevention, and treatment of cell proliferative, autoimmune/inflammatory, developmental, and neurological disorders.

### **SUMMARY OF THE INVENTION**

Various embodiments of the invention provide purified polypeptides, molecules for disease detection and treatment, referred to collectively as 'MDDT' and individually as 'MDDT-1,' 'MDDT-2,' 'MDDT-3,' 'MDDT-4,' 'MDDT-5,' 'MDDT-6,' 'MDDT-7,' 'MDDT-8,' 'MDDT-9,' 'MDDT-10,' 'MDDT-11,' 'MDDT-12,' 'MDDT-13,' 'MDDT-14,' 'MDDT-15,' 'MDDT-16,' 'MDDT-17,' 'MDDT-18,' 'MDDT-19,' 'MDDT-20,' 'MDDT-21,' 'MDDT-22,' 'MDDT-23,' 'MDDT-24,' 'MDDT-25,' 'MDDT-26,' 'MDDT-27,' 'MDDT-28,' 'MDDT-29,' 'MDDT-30,' 'MDDT-31,' 'MDDT-32,' 'MDDT-33,' 'MDDT-34,' 'MDDT-35,' 'MDDT-36,' 'MDDT-37,' 'MDDT-38,' 'MDDT-39,' 'MDDT-40,' 'MDDT-41,' 'MDDT-42,' 'MDDT-43,' 'MDDT-44,' 'MDDT-45,' 'MDDT-46,' 'MDDT-47,' 'MDDT-48,' 'MDDT-49,' 'MDDT-50,' 'MDDT-51,' 'MDDT-52,' 'MDDT-53,' 'MDDT-54,' 'MDDT-55,' 'MDDT-56,' 'MDDT-57,' 'MDDT-58,' 'MDDT-59,' 'MDDT-60,' 'MDDT-61,' 'MDDT-62,' 'MDDT-63,' 'MDDT-64,' 'MDDT-65,' 'MDDT-66,' 'MDDT-67,' 'MDDT-68,' and 'MDDT-69' and methods for using these proteins and their encoding polynucleotides for the detection, diagnosis, and treatment of diseases and medical conditions. Embodiments also provide methods for utilizing the purified molecules for disease detection and treatment and/or their encoding polynucleotides for facilitating the drug discovery process, including determination of efficacy, dosage, toxicity, and pharmacology. Related embodiments provide methods for utilizing the purified molecules for disease detection and treatment and/or their encoding polynucleotides for investigating the pathogenesis of diseases and medical conditions.

An embodiment provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-



69, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69. Another embodiment provides an isolated polypeptide comprising an amino acid sequence of SEQ ID NO:1-69.

Still another embodiment provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69. In another embodiment, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-69. In an alternative embodiment, the polynucleotide is selected from the group consisting of SEQ ID NO:70-138.

Still another embodiment provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69. Another embodiment provides a cell transformed with the recombinant polynucleotide. Yet another embodiment provides a transgenic organism comprising the recombinant polynucleotide.

Another embodiment provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69. The method comprises a) culturing a cell under conditions suitable for expression of the

polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Yet another embodiment provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69.

Still yet another embodiment provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:70-138, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:70-138, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In other embodiments, the polynucleotide can comprise at least about 20, 30, 40, 60, 80, or 100 contiguous nucleotides.

Yet another embodiment provides a method for detecting a target polynucleotide in a sample, said target polynucleotide being selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:70-138, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:70-138, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex. In a related embodiment, the method can include detecting the amount of the hybridization complex. In still other embodiments, the probe can comprise at least about 20, 30, 40, 60, 80, or 100 contiguous nucleotides.

Still yet another embodiment provides a method for detecting a target polynucleotide in a sample, said target polynucleotide being selected from the group consisting of a) a polynucleotide

comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:70-138, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:70-138, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide  
5 complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof. In a related embodiment, the method can include detecting the amount of the amplified target polynucleotide or fragment thereof.

10 Another embodiment provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, c) a biologically active  
15 fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, and a pharmaceutically acceptable excipient. In one embodiment, the composition can comprise an amino acid sequence selected from the group consisting of SEQ ID NO:1-69. Other embodiments provide a method of treating a disease or  
20 condition associated with decreased or abnormal expression of functional MDDT, comprising administering to a patient in need of such treatment the composition.

Yet another embodiment provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, b) a polypeptide comprising a  
25 naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69. The method comprises a) exposing a sample  
30 comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. Another embodiment provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. Yet another embodiment provides a method of treating a disease or condition associated with decreased expression of functional MDDT, comprising administering to a patient in need of such treatment the composition.

35 Still yet another embodiment provides a method for screening a compound for effectiveness

as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, c) a  
5 biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. Another embodiment provides a composition comprising an antagonist compound  
10 identified by the method and a pharmaceutically acceptable excipient. Yet another embodiment provides a method of treating a disease or condition associated with overexpression of functional MDDT, comprising administering to a patient in need of such treatment the composition.

Another embodiment provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid  
15 sequence selected from the group consisting of SEQ ID NO:1-69, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, and d) an immunogenic fragment of a polypeptide having an amino acid sequence  
20 selected from the group consisting of SEQ ID NO:1-69. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

Yet another embodiment provides a method of screening for a compound that modulates the  
25 activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical or at least about 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ  
30 ID NO:1-69, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the  
35 polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide

in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

Still yet another embodiment provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:70-138, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, b) detecting altered expression of the target polynucleotide, and c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

Another embodiment provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:70-138, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:70-138, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:70-138, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical or at least about 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:70-138, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide can comprise a fragment of a polynucleotide selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

### BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for full length polynucleotide and polypeptide embodiments of the invention.

Table 2 shows the GenBank identification number and annotation of the nearest GenBank

homolog, and the PROTEOME database identification numbers and annotations of PROTEOME database homologs, for polypeptide embodiments of the invention. The probability scores for the matches between each polypeptide and its homolog(s) are also shown.

Table 3 shows structural features of polypeptide embodiments, including predicted motifs  
5 and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.

Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble polynucleotide embodiments, along with selected fragments of the polynucleotides.

Table 5 shows representative cDNA libraries for polynucleotide embodiments.

10 Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze polynucleotides and polypeptides, along with applicable descriptions, references, and threshold parameters.

Table 8 shows single nucleotide polymorphisms found in polynucleotide sequences of the  
15 invention, along with allele frequencies in different human populations.

## DESCRIPTION OF THE INVENTION

Before the present proteins, nucleic acids, and methods are described, it is understood that embodiments of the invention are not limited to the particular machines, instruments, materials, and  
20 methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the invention.

As used herein and in the appended claims, the singular forms “a,” “an,” and “the” include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to “a  
25 host cell” includes a plurality of such host cells, and a reference to “an antibody” is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be  
30 used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with various embodiments of the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior  
35 invention.

**DEFINITIONS**

“MDDT” refers to the amino acid sequences of substantially purified MDDT obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

5       The term “agonist” refers to a molecule which intensifies or mimics the biological activity of MDDT. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of MDDT either by directly interacting with MDDT or by acting on components of the biological pathway in which MDDT participates.

10       An “allelic variant” is an alternative form of the gene encoding MDDT. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times  
15       in a given sequence.

      “Altered” nucleic acid sequences encoding MDDT include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as MDDT or a polypeptide with at least one functional characteristic of MDDT. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe  
20       of the polynucleotide encoding MDDT, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide encoding MDDT. The encoded protein may also be “altered,” and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent MDDT. Deliberate amino acid substitutions may be made on the basis of one or more similarities in polarity,  
25       charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of MDDT is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids  
30       with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

      The terms “amino acid” and “amino acid sequence” can refer to an oligopeptide, a peptide, a polypeptide, or a protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where “amino acid sequence” is recited to refer to a sequence of a naturally  
35       occurring protein molecule, “amino acid sequence” and like terms are not meant to limit the amino

acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

“Amplification” relates to the production of additional copies of a nucleic acid.

Amplification may be carried out using polymerase chain reaction (PCR) technologies or other nucleic acid amplification technologies well known in the art.

The term “antagonist” refers to a molecule which inhibits or attenuates the biological activity of MDDT. Antagonists may include proteins such as antibodies, anticalins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of MDDT either by directly interacting with MDDT or by acting on components of the biological pathway in which MDDT participates.

The term “antibody” refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')<sub>2</sub>, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind MDDT polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term “antigenic determinant” refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term “aptamer” refers to a nucleic acid or oligonucleotide molecule that binds to a specific molecular target. Aptamers are derived from an *in vitro* evolutionary process (e.g., SELEX (Systematic Evolution of Ligands by EXponential Enrichment), described in U.S. Patent No. 5,270,163), which selects for target-specific aptamer sequences from large combinatorial libraries.

Aptamer compositions may be double-stranded or single-stranded, and may include deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The nucleotide components of an aptamer may have modified sugar groups (e.g., the 2'-OH group of a ribonucleotide may be replaced by 2'-F or 2'-NH<sub>2</sub>), which may improve a desired property, e.g., resistance to nucleases or longer lifetime in blood. Aptamers may be conjugated to other molecules, e.g., a high molecular weight carrier to slow clearance of the aptamer from the circulatory system.



Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a cross-linker (Brody, E.N. and L. Gold (2000) J. Biotechnol. 74:5-13).

The term “intramer” refers to an aptamer which is expressed *in vivo*. For example, a vaccinia virus-based RNA expression system has been used to express specific RNA aptamers at high levels in the cytoplasm of leukocytes (Blind, M. et al. (1999) Proc. Natl. Acad. Sci. USA 96:3606-3610).

The term “spiegelmer” refers to an aptamer which includes L-DNA, L-RNA, or other left-handed nucleotide derivatives or nucleotide-like molecules. Aptamers containing left-handed nucleotides are resistant to degradation by naturally occurring enzymes, which normally act on substrates containing right-handed nucleotides.

The term “antisense” refers to any composition capable of base-pairing with the “sense” (coding) strand of a polynucleotide having a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation “negative” or “minus” can refer to the antisense strand, and the designation “positive” or “plus” can refer to the sense strand of a reference DNA molecule.

The term “biologically active” refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, “immunologically active” or “immunogenic” refers to the capability of the natural, recombinant, or synthetic MDDT, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

“Complementary” describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A “composition comprising a given polynucleotide” and a “composition comprising a given polypeptide” can refer to any composition containing the given polynucleotide or polypeptide. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotides encoding MDDT or fragments of MDDT may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's

solution, dry milk, salmon sperm DNA, etc.).

“Consensus sequence” refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (Accelrys, Burlington MA) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

“Conservative amino acid substitutions” are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
15	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
20	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
25	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
30	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
	Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A “deletion” refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term “derivative” refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an

alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

5 A “detectable label” refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

“Differential expression” refers to increased or upregulated; or decreased, downregulated, or absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a  
10 diseased and a normal sample.

“Exon shuffling” refers to the recombination of different coding regions (exons). Since an exon may represent a structural or functional domain of the encoded protein, new proteins may be assembled through the novel reassortment of stable substructures, thus allowing acceleration of the evolution of new protein functions.

15 A “fragment” is a unique portion of MDDT or a polynucleotide encoding MDDT which can be identical in sequence to, but shorter in length than, the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from about 5 to about 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes,  
20 may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any  
25 length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:70-138 can comprise a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:70-138, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:70-138 can be  
30 employed in one or more embodiments of methods of the invention, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:70-138 from related polynucleotides. The precise length of a fragment of SEQ ID NO:70-138 and the region of SEQ ID NO:70-138 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

35 A fragment of SEQ ID NO:1-69 is encoded by a fragment of SEQ ID NO:70-138. A

fragment of SEQ ID NO:1-69 can comprise a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-69. For example, a fragment of SEQ ID NO:1-69 can be used as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-69. The precise length of a fragment of SEQ ID NO:1-69 and the region of SEQ ID NO:1-69 to which the fragment corresponds can be determined based on the intended purpose for the fragment using one or more analytical methods described herein or otherwise known in the art.

A “full length” polynucleotide is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A “full length” polynucleotide sequence encodes a “full length” polypeptide sequence.

“Homology” refers to sequence similarity or, alternatively, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms “percent identity” and “% identity,” as applied to polynucleotide sequences, refer to the percentage of identical residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using one or more computer algorithms or programs known in the art or described herein. For example, percent identity can be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989; CABIOS 5:151-153) and in Higgins, D.G. et al. (1992; CABIOS 8:189-191). For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and “diagonals saved”=4. The “weighted” residue weight table is selected as the default.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms which can be used is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including “blastn,” that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called “BLAST 2 Sequences” that is used for direct pairwise comparison of two nucleotide sequences. “BLAST 2 Sequences” can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The “BLAST 2 Sequences” tool can be used for both blastn and blastp (discussed below). BLAST

programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the “BLAST 2 Sequences” tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

5 *Reward for match: 1*

*Penalty for mismatch: -2*

*Open Gap: 5 and Extension Gap: 2 penalties*

*Gap x drop-off: 50*

*Expect: 10*

10 *Word Size: 11*

*Filter: on*

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases “percent identity” and “% identity,” as applied to polypeptide sequences, refer to the percentage of identical residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide. The phrases “percent similarity” and “% similarity,” as applied to polypeptide sequences, refer to the percentage of residue matches, including identical residue matches and conservative substitutions, between at least two polypeptide sequences aligned using a standardized algorithm. In contrast, conservative substitutions are not included in the calculation of percent identity between polypeptide sequences.

Percent identity between polypeptide sequences may be determined using the default

parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62  
 Open Gap: 11 and Extension Gap: 1 penalties  
 Gap x drop-off: 50  
 Expect: 10  
 Word Size: 3  
 Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less

non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore  
5 hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about  
10 5°C to 20°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating  $T_m$  and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. and D.W. Russell (2001; Molecular Cloning: A Laboratory Manual, 3rd ed., vol. 1-3, Cold Spring Harbor Press,  
15 Cold Spring Harbor NY, ch. 9).

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%.  
20 Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high  
25 stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term “hybridization complex” refers to a complex formed between two nucleic acids by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g.,  $C_0t$  or  $R_0t$  analysis) or formed between one nucleic acid present in  
30 solution and another nucleic acid immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words “insertion” and “addition” refer to changes in an amino acid or polynucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

“Immune response” can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

5       An “immunogenic fragment” is a polypeptide or oligopeptide fragment of MDDT which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term “immunogenic fragment” also includes any polypeptide or oligopeptide fragment of MDDT which is useful in any of the antibody production methods disclosed herein or known in the art.

10       The term “microarray” refers to an arrangement of a plurality of polynucleotides, polypeptides, antibodies, or other chemical compounds on a substrate.

The terms “element” and “array element” refer to a polynucleotide, polypeptide, antibody, or other chemical compound having a unique and defined position on a microarray.

15       The term “modulate” refers to a change in the activity of MDDT. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of MDDT.

20       The phrases “nucleic acid” and “nucleic acid sequence” refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

25       “Operably linked” refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

30       “Peptide nucleic acid” (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

“Post-translational modification” of an MDDT may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of MDDT.



“Probe” refers to nucleic acids encoding MDDT, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acids. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. “Primers” are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in, for example, Sambrook, J. and D.W. Russell (2001; Molecular Cloning: A Laboratory Manual, 3rd ed., vol. 1-3, Cold Spring Harbor Press, Cold Spring Harbor NY), Ausubel, F.M. et al. (1999; Short Protocols in Molecular Biology, 4<sup>th</sup> ed., John Wiley & Sons, New York NY), and Innis, M. et al. (1990; PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA). PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a “mispriming library,” in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection

programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a nucleic acid that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook and Russell (*supra*). The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA molecule, is composed of the same linear sequence of nucleotides as the reference DNA molecule with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing MDDT, nucleic acids encoding MDDT, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

5       The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide  
10       comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

      The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably at least about 75% free, and most preferably at least about 90% free from other  
15       components with which they are naturally associated.

      A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

      "Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers,  
20       microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

      A "transcript image" or "expression profile" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

      "Transformation" describes a process by which exogenous DNA is introduced into a recipient  
25       cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term  
30       "transformed cells" includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

      A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic

acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. In another embodiment, the nucleic acid can be introduced by infection with a recombinant viral vector, such as a lentiviral vector (Lois, C. et al. (2002) Science 295:868-872). The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook and Russell (*supra*).

A “variant” of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the “BLAST 2 Sequences” tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an “allelic” (as defined above), “splice,” “species,” or “polymorphic” variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotides that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass “single nucleotide polymorphisms” (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A “variant” of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity or sequence similarity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the “BLAST 2 Sequences” tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for

example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity or sequence similarity over a certain defined length of one of the polypeptides.

5

## THE INVENTION

Various embodiments of the invention include new human molecules for disease detection and treatment (MDDT), the polynucleotides encoding MDDT, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, autoimmune/inflammatory,

10 developmental, and neurological disorders.

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide embodiments of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an  
15 Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown.

Table 2 shows sequences with homology to polypeptide embodiments of the invention as  
20 identified by BLAST analysis against the GenBank protein (genpept) database and the PROTEOME database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (GenBank ID NO:) of the nearest GenBank homolog and the PROTEOME database identification numbers  
25 (PROTEOME ID NO:) of the nearest PROTEOME database homologs. Column 4 shows the probability scores for the matches between each polypeptide and its homolog(s). Column 5 shows the annotation of the GenBank and PROTEOME database homolog(s) along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of the polypeptides of the invention. Columns 1  
30 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Accelrys, Burlington MA).

Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are molecules for disease detection and treatment. For example, SEQ ID NO:5 has homology to a protein which appears to localize to membranes, as determined by BLAST analysis using the PROTEOME database (PROTEOME ID 370403|SPBC20F10.07). SEQ ID NO:5 also contains a GRAM domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from TMHMMER, MOTIFS, and additional BLAST analyses provide further corroborative evidence that SEQ ID NO:5 is a membrane-associated protein. In an alternative example, SEQ ID NO:24 contains a SET domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) In an alternative example, SEQ ID NO:55 is 38% identical, from residue K16 to residue W298, to Podospira anserina beta transducin-line protein (GenBank ID g607003) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is  $1.3e-47$ , which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:55 also has homology to proteins that contain WD domains, as determined by BLAST analysis using the PROTEOME database. SEQ ID NO:55 also contains a WD40 repeat domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS and MOTIFS analyses provide further corroborative evidence that SEQ ID NO:55 is a WD repeat protein. In an alternative example, SEQ ID NO:68 is 61% identical, from residue G62 to residue K563, to mouse DMR-N9 (GenBank ID g817954) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is  $8.2e-173$ , which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:68 also has homology to proteins that are associated with myotonic dystrophy, as determined by BLAST analysis using the PROTEOME database. SEQ ID NO:68 also contains WD repeats as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM and SMART databases of conserved protein families/domains. (See Table 3.) Data from BLIMPS and BLAST analyses provide further corroborative evidence that SEQ ID NO:68 is a G-beta WD repeat protein. SEQ ID NO:1-4, SEQ ID NO:6-23, SEQ ID NO:25-54, SEQ ID NO:56-67, and SEQ ID NO:69 were analyzed and annotated in a similar manner. The algorithms and parameters for the

analysis of SEQ ID NO:1-69 are described in Table 7.

As shown in Table 4, the full length polynucleotide embodiments were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Column 1 lists the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:), the corresponding Incyte polynucleotide consensus sequence number (Incyte ID) for each polynucleotide of the invention, and the length of each polynucleotide sequence in basepairs. Column 2 shows the nucleotide start (5') and stop (3') positions of the cDNA and/or genomic sequences used to assemble the full length polynucleotide embodiments, and of fragments of the polynucleotides which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:70-138 or that distinguish between SEQ ID NO:70-138 and related polynucleotides.

The polynucleotide fragments described in Column 2 of Table 4 may refer specifically, for example, to Incyte cDNAs derived from tissue-specific cDNA libraries or from pooled cDNA libraries. Alternatively, the polynucleotide fragments described in column 2 may refer to GenBank cDNAs or ESTs which contributed to the assembly of the full length polynucleotides. In addition, the polynucleotide fragments described in column 2 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the polynucleotide fragments described in column 2 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the polynucleotide fragments described in column 2 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, a polynucleotide sequence identified as FL\_XXXXXX\_N<sub>1</sub>\_N<sub>2</sub>\_YYYYY\_N<sub>3</sub>\_N<sub>4</sub> represents a "stitched" sequence in which XXXXXX is the identification number of the cluster of sequences to which the algorithm was applied, and YYYYY is the number of the prediction generated by the algorithm, and N<sub>1,2,3...</sub>, if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the polynucleotide fragments in column 2 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, a polynucleotide sequence identified as FLXXXXXX\_gAAAAA\_gBBBBB\_1\_N is a "stretched" sequence, with XXXXXX being the Incyte project identification number, gAAAAA being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, gBBBBB being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and N referring to specific exons (See Example V). In instances where a RefSeq

sequence was used as a protein homolog for the “exon-stretching” algorithm, a RefSeq identifier (denoted by “NM,” “NP,” or “NT”) may be used in place of the GenBank identifier (*i.e.*, *gBBBBB*).

Alternatively, a prefix identifies component sequences that were hand-edited, predicted from genomic DNA sequences, or derived from a combination of sequence analysis methods. The

5 following Table lists examples of component sequence prefixes and corresponding sequence analysis methods associated with the prefixes (see Example IV and Example V).

Prefix	Type of analysis and/or examples of programs
GNN, GFG, ENST	Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK).
GBI	Hand-edited analysis of genomic sequences.
FL	Stitched or stretched genomic sequences (see Example V).
INCY	Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript.

10 In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in Table 4 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

Table 5 shows the representative cDNA libraries for those full length polynucleotides which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotides. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

Table 8 shows single nucleotide polymorphisms (SNPs) found in polynucleotide sequences of the invention, along with allele frequencies in different human populations. Columns 1 and 2 show the polynucleotide sequence identification number (SEQ ID NO:) and the corresponding Incyte project identification number (PID) for polynucleotides of the invention. Column 3 shows the Incyte identification number for the EST in which the SNP was detected (EST ID), and column 4 shows the identification number for the SNP (SNP ID). Column 5 shows the position within the EST sequence at which the SNP is located (EST SNP), and column 6 shows the position of the SNP within the full-length polynucleotide sequence (CB1 SNP). Column 7 shows the allele found in the EST sequence. Columns 8 and 9 show the two alleles found at the SNP site. Column 10 shows the amino acid encoded by the codon including the SNP site, based upon the allele found in the EST. Columns 11-



14 show the frequency of allele 1 in four different human populations. An entry of n/d (not detected) indicates that the frequency of allele 1 in the population was too low to be detected, while n/a (not available) indicates that the allele frequency was not determined for the population.

The invention also encompasses MDDT variants. Various embodiments of MDDT variants  
5 can have at least about 80%, at least about 90%, or at least about 95% amino acid sequence identity to the MDDT amino acid sequence, and can contain at least one functional or structural characteristic of MDDT.

Various embodiments also encompass polynucleotides which encode MDDT. In a particular  
10 embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:70-138, which encodes MDDT. The polynucleotide sequences of SEQ ID NO:70-138, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses variants of a polynucleotide encoding MDDT. In particular,  
15 such a variant polynucleotide will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a polynucleotide encoding MDDT. A particular aspect of the invention encompasses a variant of a polynucleotide comprising a sequence selected from the group consisting of SEQ ID NO:70-138 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a  
20 nucleic acid sequence selected from the group consisting of SEQ ID NO:70-138. Any one of the polynucleotide variants described above can encode a polypeptide which contains at least one functional or structural characteristic of MDDT.

In addition, or in the alternative, a polynucleotide variant of the invention is a splice variant of a polynucleotide encoding MDDT. A splice variant may have portions which have significant  
25 sequence identity to a polynucleotide encoding MDDT, but will generally have a greater or lesser number of polynucleotides due to additions or deletions of blocks of sequence arising from alternate splicing of exons during mRNA processing. A splice variant may have less than about 70%, or alternatively less than about 60%, or alternatively less than about 50% polynucleotide sequence identity to a polynucleotide encoding MDDT over its entire length; however, portions of the splice  
30 variant will have at least about 70%, or alternatively at least about 85%, or alternatively at least about 95%, or alternatively 100% polynucleotide sequence identity to portions of the polynucleotide encoding MDDT. For example, a polynucleotide comprising a sequence of SEQ ID NO:77, a polynucleotide comprising a sequence of SEQ ID NO:81, and a polynucleotide comprising a sequence of SEQ ID NO:95 are splice variants of each other; a polynucleotide comprising a sequence

of SEQ ID NO:83 and a polynucleotide comprising a sequence of SEQ ID NO:93 are splice variants of each other; a polynucleotide comprising a sequence of SEQ ID NO:85 and a polynucleotide comprising a sequence of SEQ ID NO:92 are splice variants of each other; a polynucleotide comprising a sequence of SEQ ID NO:115 and a polynucleotide comprising a sequence of SEQ ID NO:121 are splice variants of each other; a polynucleotide comprising a sequence of SEQ ID NO:131 and a polynucleotide comprising a sequence of SEQ ID NO:134 are splice variants of each other; and a polynucleotide comprising a sequence of SEQ ID NO:132 and a polynucleotide comprising a sequence of SEQ ID NO:133 are splice variants of each other. Any one of the splice variants described above can encode a polypeptide which contains at least one functional or structural characteristic of MDDT.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding MDDT, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring MDDT, and all such variations are to be considered as being specifically disclosed.

Although polynucleotides which encode MDDT and its variants are generally capable of hybridizing to polynucleotides encoding naturally occurring MDDT under appropriately selected conditions of stringency, it may be advantageous to produce polynucleotides encoding MDDT or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding MDDT and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of polynucleotides which encode MDDT and MDDT derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic polynucleotide may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a polynucleotide encoding MDDT or any fragment thereof.

Embodiments of the invention can also include polynucleotides that are capable of

hybridizing to the claimed polynucleotides, and, in particular, to those having the sequences shown in SEQ ID NO:70-138 and fragments thereof, under various conditions of stringency (Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.*

152:507-511). Hybridization conditions, including annealing and wash conditions, are described in

5 "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Biosciences, Piscataway NJ), or combinations  
10 of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Invitrogen, Carlsbad CA). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems).

Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied  
15 Biosystems), the MEGABACE 1000 DNA sequencing system (Amersham Biosciences), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art (Ausubel et al., *supra*, ch. 7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853).

The nucleic acids encoding MDDT may be extended utilizing a partial nucleotide sequence  
20 and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector (Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322). Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a  
25 circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences (Triglia, T. et al. (1988) *Nucleic Acids Res.* 16:8186). A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA (Lagerstrom, M. et al. (1991) *PCR Methods Applic.* 1:111-119). In this method, multiple restriction enzyme digestions and ligations  
30 may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art (Parker, J.D. et al. (1991) *Nucleic Acids Res.* 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon

junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

5           When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

10           Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate  
15           software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

            In another embodiment of the invention, polynucleotides or fragments thereof which encode  
20           MDDT may be cloned in recombinant DNA molecules that direct expression of MDDT, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other polynucleotides which encode substantially the same or a functionally equivalent polypeptides may be produced and used to express MDDT.

            The polynucleotides of the invention can be engineered using methods generally known in  
25           the art in order to alter MDDT-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction  
30           sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

            The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent No. 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or

improve the biological properties of MDDT, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired  
5 properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of  
10 homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, polynucleotides encoding MDDT may be synthesized, in whole or in part, using one or more chemical methods well known in the art (Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232).  
15 Alternatively, MDDT itself or a fragment thereof may be synthesized using chemical methods known in the art. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques (Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; Roberge, J.Y. et al. (1995) Science 269:202-204). Automated  
20 synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of MDDT, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid  
25 chromatography (Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421). The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing (Creighton, *supra*, pp. 28-53).

In order to express a biologically active MDDT, the polynucleotides encoding MDDT or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains  
30 the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotides encoding MDDT. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of polynucleotides encoding MDDT. Such signals

include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where a polynucleotide sequence encoding MDDT and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used (Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162).

Methods which are well known to those skilled in the art may be used to construct expression vectors containing polynucleotides encoding MDDT and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination (Sambrook and Russell, *supra*, ch. 1-4, and 8; Ausubel et al., *supra*, ch. 1, 3, and 15).

A variety of expression vector/host systems may be utilized to contain and express polynucleotides encoding MDDT. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems (Sambrook and Russell, *supra*; Ausubel et al., *supra*; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355). Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of polynucleotides to the targeted organ, tissue, or cell population (Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5:350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90:6340-6344; Buller, R.M. et al. (1985) Nature 317:813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31:219-226; Verma, I.M. and N. Somia (1997) Nature 389:239-242). The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotides encoding MDDT. For example, routine cloning,

subcloning, and propagation of polynucleotides encoding MDDT can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSORT1 plasmid (Invitrogen). Ligation of polynucleotides encoding MDDT into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence (Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509). When large quantities of MDDT are needed, e.g. for the production of antibodies, vectors which direct high level expression of MDDT may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of MDDT. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign polynucleotide sequences into the host genome for stable propagation (Ausubel et al., *supra*; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; Scorer, C.A. et al. (1994) Bio/Technology 12:181-184).

Plant systems may also be used for expression of MDDT. Transcription of polynucleotides encoding MDDT may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection (The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196).

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, polynucleotides encoding MDDT may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses MDDT in host cells (Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes (Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355).

5 For long term production of recombinant proteins in mammalian systems, stable expression of MDDT in cell lines is preferred. For example, polynucleotides encoding MDDT can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days  
10 in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These  
15 include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk<sup>-</sup>* and *apr<sup>-</sup>* cells, respectively (Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823). Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat*  
20 confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14). Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites (Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051). Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech),  $\beta$ -  
25 glucuronidase and its substrate  $\beta$ -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is  
30 also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding MDDT is inserted within a marker gene sequence, transformed cells containing polynucleotides encoding MDDT can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding MDDT under the control of a single promoter. Expression of the marker gene in response to induction or selection



usually indicates expression of the tandem gene as well.

In general, host cells that contain the polynucleotide encoding MDDT and that express MDDT may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR  
5 amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of MDDT using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and  
10 fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on MDDT is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art (Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-  
15 Interscience, New York NY; Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding MDDT  
20 include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, polynucleotides encoding MDDT, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted  
25 using a variety of commercially available kits, such as those provided by Amersham Biosciences, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with polynucleotides encoding MDDT may be cultured under  
30 conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode MDDT may be designed to contain signal sequences which direct secretion of MDDT through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted polynucleotides or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a “prepro” or “pro” form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant polynucleotides encoding MDDT may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric MDDT protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of MDDT activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the MDDT encoding sequence and the heterologous protein sequence, so that MDDT may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel et al. (*supra*, ch. 10 and 16). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In another embodiment, synthesis of radiolabeled MDDT may be achieved *in vitro* using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, <sup>35</sup>S-methionine.

MDDT, fragments of MDDT, or variants of MDDT may be used to screen for compounds that specifically bind to MDDT. One or more test compounds may be screened for specific binding

to MDDT. In various embodiments, 1, 2, 3, 4, 5, 10, 20, 50, 100, or 200 test compounds can be screened for specific binding to MDDT. Examples of test compounds can include antibodies, anticalins, oligonucleotides, proteins (e.g., ligands or receptors), or small molecules.

In related embodiments, variants of MDDT can be used to screen for binding of test compounds, such as antibodies, to MDDT, a variant of MDDT, or a combination of MDDT and/or one or more variants MDDT. In an embodiment, a variant of MDDT can be used to screen for compounds that bind to a variant of MDDT, but not to MDDT having the exact sequence of a sequence of SEQ ID NO:1-69. MDDT variants used to perform such screening can have a range of about 50% to about 99% sequence identity to MDDT, with various embodiments having 60%, 70%, 75%, 80%, 85%, 90%, and 95% sequence identity.

In an embodiment, a compound identified in a screen for specific binding to MDDT can be closely related to the natural ligand of MDDT, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner (Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2):Chapter 5). In another embodiment, the compound thus identified can be a natural ligand of a receptor MDDT (Howard, A.D. et al. (2001) Trends Pharmacol. Sci.22:132-140; Wise, A. et al. (2002) Drug Discovery Today 7:235-246).

In other embodiments, a compound identified in a screen for specific binding to MDDT can be closely related to the natural receptor to which MDDT binds, at least a fragment of the receptor, or a fragment of the receptor including all or a portion of the ligand binding site or binding pocket. For example, the compound may be a receptor for MDDT which is capable of propagating a signal, or a decoy receptor for MDDT which is not capable of propagating a signal (Ashkenazi, A. and V.M. Divit (1999) Curr. Opin. Cell Biol. 11:255-260; Mantovani, A. et al. (2001) Trends Immunol. 22:328-336). The compound can be rationally designed using known techniques. Examples of such techniques include those used to construct the compound etanercept (ENBREL; Amgen Inc., Thousand Oaks CA), which is efficacious for treating rheumatoid arthritis in humans. Etanercept is an engineered p75 tumor necrosis factor (TNF) receptor dimer linked to the Fc portion of human IgG<sub>1</sub> (Taylor, P.C. et al. (2001) Curr. Opin. Immunol. 13:611-616).

In one embodiment, two or more antibodies having similar or, alternatively, different specificities can be screened for specific binding to MDDT, fragments of MDDT, or variants of MDDT. The binding specificity of the antibodies thus screened can thereby be selected to identify particular fragments or variants of MDDT. In one embodiment, an antibody can be selected such that its binding specificity allows for preferential identification of specific fragments or variants of MDDT. In another embodiment, an antibody can be selected such that its binding specificity allows for preferential diagnosis of a specific disease or condition having increased, decreased, or otherwise

abnormal production of MDDT.

In an embodiment, anticalins can be screened for specific binding to MDDT, fragments of MDDT, or variants of MDDT. Anticalins are ligand-binding proteins that have been constructed based on a lipocalin scaffold (Weiss, G.A. and H.B. Lowman (2000) Chem. Biol. 7:R177-R184;

5 Skerra, A. (2001) J. Biotechnol. 74:257-275). The protein architecture of lipocalins can include a beta-barrel having eight antiparallel beta-strands, which supports four loops at its open end. These loops form the natural ligand-binding site of the lipocalins, a site which can be re-engineered *in vitro* by amino acid substitutions to impart novel binding specificities. The amino acid substitutions can be made using methods known in the art or described herein, and can include conservative substitutions  
10 (e.g., substitutions that do not alter binding specificity) or substitutions that modestly, moderately, or significantly alter binding specificity.

In one embodiment, screening for compounds which specifically bind to, stimulate, or inhibit MDDT involves producing appropriate cells which express MDDT, either as a secreted protein or on the cell membrane. Preferred cells can include cells from mammals, yeast, *Drosophila*, or *E. coli*.

15 Cells expressing MDDT or cell membrane fractions which contain MDDT are then contacted with a test compound and binding, stimulation, or inhibition of activity of either MDDT or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example,  
20 the assay may comprise the steps of combining at least one test compound with MDDT, either in solution or affixed to a solid support, and detecting the binding of MDDT to the compound.

Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a  
25 solid support.

An assay can be used to assess the ability of a compound to bind to its natural ligand and/or to inhibit the binding of its natural ligand to its natural receptors. Examples of such assays include radio-labeling assays such as those described in U.S. Patent No. 5,914,236 and U.S. Patent No. 6,372,724. In a related embodiment, one or more amino acid substitutions can be introduced into a  
30 polypeptide compound (such as a receptor) to improve or alter its ability to bind to its natural ligands (Matthews, D.J. and J.A. Wells. (1994) Chem. Biol. 1:25-30). In another related embodiment, one or more amino acid substitutions can be introduced into a polypeptide compound (such as a ligand) to improve or alter its ability to bind to its natural receptors (Cunningham, B.C. and J.A. Wells (1991) Proc. Natl. Acad. Sci. USA 88:3407-3411; Lowman, H.B. et al. (1991) J. Biol. Chem. 266:10982-

10988).

MDDT, fragments of MDDT, or variants of MDDT may be used to screen for compounds that modulate the activity of MDDT. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for MDDT activity, wherein MDDT is combined with at least one test compound, and the activity of MDDT in the presence of a test compound is compared with the activity of MDDT in the absence of the test compound. A change in the activity of MDDT in the presence of the test compound is indicative of a compound that modulates the activity of MDDT. Alternatively, a test compound is combined with an *in vitro* or cell-free system comprising MDDT under conditions suitable for MDDT activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of MDDT may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding MDDT or their mammalian homologs may be “knocked out” in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease (see, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337). For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (*neo*; Capecchi, M.R. (1989) *Science* 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) *Clin. Invest.* 97:1999-2002; Wagner, K.U. et al. (1997) *Nucleic Acids Res.* 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding MDDT may also be manipulated *in vitro* in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) *Science* 282:1145-1147).

Polynucleotides encoding MDDT can also be used to create “knockin” humanized animals

(pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding MDDT is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress MDDT, e.g., by secreting MDDT in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) *Biotechnol. Annu. Rev.* 4:55-74).

## THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of MDDT and molecules for disease detection and treatment. In addition, examples of tissues expressing MDDT can be found in Table 6 and can also be found in Example XI. Therefore, MDDT appears to play a role in cell proliferative, autoimmune/inflammatory, developmental, and neurological disorders. In the treatment of disorders associated with increased MDDT expression or activity, it is desirable to decrease the expression or activity of MDDT. In the treatment of disorders associated with decreased MDDT expression or activity, it is desirable to increase the expression or activity of MDDT.

Therefore, in one embodiment, MDDT or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MDDT. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, colon, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or

pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, 5 fungal, parasitic, protozoal, and helminthic infections, and trauma; a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such 10 as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; and a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal 15 disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann- 20 Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other 25 neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, 30 and familial frontotemporal dementia.

In another embodiment, a vector capable of expressing MDDT or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MDDT including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified MDDT in

conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MDDT including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of MDDT may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MDDT including, but not limited to, those listed above.

In a further embodiment, an antagonist of MDDT may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of MDDT. Examples of such disorders include, but are not limited to, those cell proliferative, autoimmune/inflammatory, developmental, and neurological disorders described above. In one aspect, an antibody which specifically binds MDDT may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express MDDT.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding MDDT may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of MDDT including, but not limited to, those described above.

In other embodiments, any protein, agonist, antagonist, antibody, complementary sequence, or vector embodiments may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of MDDT may be produced using methods which are generally known in the art. In particular, purified MDDT may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind MDDT. Antibodies to MDDT may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. In an embodiment, neutralizing antibodies (i.e., those which inhibit dimer formation) can be used therapeutically. Single chain antibodies (e.g., from camels or llamas) may be potent enzyme inhibitors and may have application in the design of peptide mimetics, and in the development of immuno-adsorbents and biosensors (Muyldermans, S. (2001) J. Biotechnol. 74:277-302).

For the production of antibodies, various hosts including goats, rabbits, rats, mice, camels, dromedaries, llamas, humans, and others may be immunized by injection with MDDT or with any



fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol.

- 5 Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to MDDT have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or  
10 fragments are substantially identical to a portion of the amino acid sequence of the natural protein. Short stretches of MDDT amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to MDDT may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not  
15 limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120).

In addition, techniques developed for the production of "chimeric antibodies," such as the  
20 splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; Takeda, S. et al. (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce MDDT-specific single chain  
25 antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137).

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as  
30 disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

Antibody fragments which contain specific binding sites for MDDT may also be generated. For example, such fragments include, but are not limited to, F(ab')<sub>2</sub> fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of

the F(ab')<sub>2</sub> fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W.D. et al. (1989) Science 246:1275-1281).

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between MDDT and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering MDDT epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for MDDT. Affinity is expressed as an association constant,  $K_a$ , which is defined as the molar concentration of MDDT-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The  $K_a$  determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple MDDT epitopes, represents the average affinity, or avidity, of the antibodies for MDDT. The  $K_a$  determined for a preparation of monoclonal antibodies, which are monospecific for a particular MDDT epitope, represents a true measure of affinity. High-affinity antibody preparations with  $K_a$  ranging from about  $10^9$  to  $10^{12}$  L/mole are preferred for use in immunoassays in which the MDDT-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with  $K_a$  ranging from about  $10^6$  to  $10^7$  L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of MDDT, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of MDDT-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available (Catty, *supra*; Coligan et al., *supra*).

In another embodiment of the invention, polynucleotides encoding MDDT, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene

expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding MDDT. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding

5 MDDT (Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press, Totawa NJ).

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein (Slater, J.E. et

10 al. (1998) J. Allergy Clin. Immunol. 102:469-475; Scanlon, K.J. et al. (1995) 9:1288-1296).

Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors (Miller, A.D. (1990) Blood 76:271; Ausubel et al., *supra*; Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63:323-347). Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in

15 the art (Rossi, J.J. (1995) Br. Med. Bull. 51:217-225; Boado, R.J. et al. (1998) J. Pharm. Sci. 87:1308-1315; Morris, M.C. et al. (1997) Nucleic Acids Res. 25:2730-2736).

In another embodiment of the invention, polynucleotides encoding MDDT may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-

20 linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassamias, familial

25 hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and N. Somia (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D.

30 (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as *Candida albicans* and *Paracoccidioides brasiliensis*; and protozoan parasites such as *Plasmodium falciparum* and *Trypanosoma cruzi*). In the case where a genetic deficiency in MDDT expression or regulation causes disease, the expression of MDDT from an appropriate population of transduced cells may alleviate the clinical manifestations

caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in MDDT are treated by constructing mammalian expression vectors encoding MDDT and introducing these vectors by mechanical means into MDDT-deficient cells. Mechanical transfer technologies for use with cells *in vivo* or *ex vitro* include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev. Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J.-L. and H. Récipon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

Expression vectors that may be effective for the expression of MDDT include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX, PCR2-TOPOTA vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). MDDT may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or  $\beta$ -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551; Gossen, M. et al. (1995) *Science* 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, *supra*), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding MDDT from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) *Virology* 52:456-467), or by electroporation (Neumann, E. et al. (1982) *EMBO J.* 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to MDDT expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding MDDT under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences

required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for  
5 receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent No. 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant")  
10 discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4<sup>+</sup> T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716;  
15 Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In an embodiment, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding MDDT to cells which have one or more genetic abnormalities with respect to the expression of MDDT. The construction and packaging of adenovirus-based vectors are well  
20 known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent No. 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999; Annu.  
25 Rev. Nutr. 19:511-544) and Verma, I.M. and N. Somia (1997; Nature 18:389:239-242).

In another embodiment, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding MDDT to target cells which have one or more genetic abnormalities with respect to the expression of MDDT. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing MDDT to cells of the central nervous system, for which HSV has  
30 a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby

incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22.

5 For HSV vectors, see also Goins, W.F. et al. (1999; J. Virol. 73:519-532) and Xu, H. et al. (1994; Dev. Biol. 163:152-161). The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

10 In another embodiment, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding MDDT to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid  
15 proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for MDDT into the alphavirus genome in place of the capsid-coding region results in the production of a large number of MDDT-coding RNAs and the synthesis of high levels of MDDT in vector transduced cells. While  
20 alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of MDDT into a variety of cell types. The specific transduction of a subset of  
25 cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions  
30 -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr,

Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177). A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of RNA molecules encoding MDDT.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA molecules encoding MDDT. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

In other embodiments of the invention, the expression of one or more selected polynucleotides of the present invention can be altered, inhibited, decreased, or silenced using RNA interference (RNAi) or post-transcriptional gene silencing (PTGS) methods known in the art. RNAi is a post-transcriptional mode of gene silencing in which double-stranded RNA (dsRNA) introduced

into a targeted cell specifically suppresses the expression of the homologous gene (i.e., the gene bearing the sequence complementary to the dsRNA). This effectively knocks out or substantially reduces the expression of the targeted gene. PTGS can also be accomplished by use of DNA or DNA fragments as well. RNAi methods are described by Fire, A. et al. (1998; Nature 391:806-811) and Gura, T. (2000; Nature 404:804-808). PTGS can also be initiated by introduction of a complementary segment of DNA into the selected tissue using gene delivery and/or viral vector delivery methods described herein or known in the art.

RNAi can be induced in mammalian cells by the use of small interfering RNA also known as siRNA. SiRNA are shorter segments of dsRNA (typically about 21 to 23 nucleotides in length) that result *in vivo* from cleavage of introduced dsRNA by the action of an endogenous ribonuclease. SiRNA appear to be the mediators of the RNAi effect in mammals. The most effective siRNAs appear to be 21 nucleotide dsRNAs with 2 nucleotide 3' overhangs. The use of siRNA for inducing RNAi in mammalian cells is described by Elbashir, S.M. et al. (2001; Nature 411:494-498).

SiRNA can either be generated indirectly by introduction of dsRNA into the targeted cell, or directly by mammalian transfection methods and agents described herein or known in the art (such as liposome-mediated transfection, viral vector methods, or other polynucleotide delivery/introductory methods). Suitable SiRNAs can be selected by examining a transcript of the target polynucleotide (e.g., mRNA) for nucleotide sequences downstream from the AUG start codon and recording the occurrence of each nucleotide and the 3' adjacent 19 to 23 nucleotides as potential siRNA target sites, with sequences having a 21 nucleotide length being preferred. Regions to be avoided for target siRNA sites include the 5' and 3' untranslated regions (UTRs) and regions near the start codon (within 75 bases), as these may be richer in regulatory protein binding sites. UTR-binding proteins and/or translation initiation complexes may interfere with binding of the siRNP endonuclease complex. The selected target sites for siRNA can then be compared to the appropriate genome database (e.g., human, etc.) using BLAST or other sequence comparison algorithms known in the art. Target sequences with significant homology to other coding sequences can be eliminated from consideration. The selected SiRNAs can be produced by chemical synthesis methods known in the art or by *in vitro* transcription using commercially available methods and kits such as the SILENCER siRNA construction kit (Ambion, Austin TX).

In alternative embodiments, long-term gene silencing and/or RNAi effects can be induced in selected tissue using expression vectors that continuously express siRNA. This can be accomplished using expression vectors that are engineered to express hairpin RNAs (shRNAs) using methods known in the art (see, e.g., Brummelkamp, T.R. et al. (2002) Science 296:550-553; and Paddison, P.J. et al. (2002) Genes Dev. 16:948-958). In these and related embodiments, shRNAs can be delivered to



target cells using expression vectors known in the art. An example of a suitable expression vector for delivery of siRNA is the PSILENCER1.0-U6 (circular) plasmid (Ambion). Once delivered to the target tissue, shRNAs are processed *in vivo* into siRNA-like molecules capable of carrying out gene-specific silencing.

5           In various embodiments, the expression levels of genes targeted by RNAi or PTGS methods can be determined by assays for mRNA and/or protein analysis. Expression levels of the mRNA of a targeted gene, can be determined by northern analysis methods using, for example, the NORTHERNMAX-GLY kit (Ambion); by microarray methods; by PCR methods; by real time PCR methods; and by other RNA/polynucleotide assays known in the art or described herein. Expression  
10       levels of the protein encoded by the targeted gene can be determined by Western analysis using standard techniques known in the art.

          An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding MDDT. Compounds which may be effective in altering expression of a specific polynucleotide may include,  
15       but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased  
20       MDDT expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding MDDT may be therapeutically useful, and in the treatment of disorders associated with decreased MDDT expression or activity, a compound which specifically promotes expression of the polynucleotide encoding MDDT may be therapeutically useful.

          In various embodiments, one or more test compounds may be screened for effectiveness in  
25       altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a  
30       library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding MDDT is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding MDDT are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is

detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding MDDT. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use *in vivo*, *in vitro*, and *ex vivo*. For *ex vivo* therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art (Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466).

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of MDDT, antibodies to MDDT, and mimetics, agonists, antagonists, or inhibitors of MDDT.

In various embodiments, the compositions described herein, such as pharmaceutical compositions, may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form.

These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery allows administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising MDDT or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, MDDT or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example MDDT or fragments thereof, antibodies of MDDT, and agonists, antagonists or inhibitors of MDDT, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the  $ED_{50}$  (the dose therapeutically effective in 50% of the population) or  $LD_{50}$  (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the  $LD_{50}/ED_{50}$  ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the  $ED_{50}$  with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1  $\mu\text{g}$  to 100,000  $\mu\text{g}$ , up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

#### DIAGNOSTICS

In another embodiment, antibodies which specifically bind MDDT may be used for the diagnosis of disorders characterized by expression of MDDT, or in assays to monitor patients being treated with MDDT or agonists, antagonists, or inhibitors of MDDT. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for MDDT include methods which utilize the antibody and a label to detect MDDT in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring MDDT, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of MDDT expression. Normal or standard values for MDDT expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to MDDT under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of MDDT expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, polynucleotides encoding MDDT may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotides, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect

and quantify gene expression in biopsied tissues in which expression of MDDT may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of MDDT, and to monitor regulation of MDDT levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotides, including genomic sequences, encoding MDDT or closely related molecules may be used to identify nucleic acid sequences which encode MDDT. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding MDDT, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the MDDT encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:70-138 or from genomic sequences including promoters, enhancers, and introns of the MDDT gene.

Means for producing specific hybridization probes for polynucleotides encoding MDDT include the cloning of polynucleotides encoding MDDT or MDDT derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as <sup>32</sup>P or <sup>35</sup>S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotides encoding MDDT may be used for the diagnosis of disorders associated with expression of MDDT. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, colon, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia

with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; and a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia. Polynucleotides encoding MDDT may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick,

pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered MDDT expression. Such qualitative or quantitative methods are well known in the art.

In a particular embodiment, polynucleotides encoding MDDT may be used in assays that detect the presence of associated disorders, particularly those mentioned above. Polynucleotides complementary to sequences encoding MDDT may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of polynucleotides encoding MDDT in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of MDDT, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding MDDT, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier, thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding MDDT may involve the use of PCR. These oligomers may be chemically synthesized, generated

enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding MDDT, or a fragment of a polynucleotide complementary to the polynucleotide encoding MDDT, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or  
5 quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from polynucleotides encoding MDDT may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP)  
10 and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from polynucleotides encoding MDDT are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel  
15 electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-  
20 based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

SNPs may be used to study the genetic basis of human disease. For example, at least 16  
25 common SNPs have been associated with non-insulin-dependent diabetes mellitus. SNPs are also useful for examining differences in disease outcomes in monogenic disorders, such as cystic fibrosis, sickle cell anemia, or chronic granulomatous disease. For example, variants in the mannose-binding lectin, MBL2, have been shown to be correlated with deleterious pulmonary outcomes in cystic fibrosis. SNPs also have utility in pharmacogenomics, the identification of genetic variants that  
30 influence a patient's response to a drug, such as life-threatening toxicity. For example, a variation in N-acetyl transferase is associated with a high incidence of peripheral neuropathy in response to the anti-tuberculosis drug isoniazid, while a variation in the core promoter of the ALOX5 gene results in diminished clinical response to treatment with an anti-asthma drug that targets the 5-lipoxygenase pathway. Analysis of the distribution of SNPs in different populations is useful for investigating



genetic drift, mutation, recombination, and selection, as well as for tracing the origins of populations and their migrations (Taylor, J.G. et al. (2001) *Trends Mol. Med.* 7:507-512; Kwok, P.-Y. and Z. Gu (1999) *Mol. Med. Today* 5:538-543; Nowotny, P. et al. (2001) *Curr. Opin. Neurobiol.* 11:637-641).

Methods which may also be used to quantify the expression of MDDT include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves (Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplaa, C. et al. (1993) *Anal. Biochem.* 212:229-236). The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotides described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, MDDT, fragments of MDDT, or antibodies specific for MDDT may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time (Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484; hereby expressly incorporated by reference herein). Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present

invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression *in vivo*,  
5 as in the case of a tissue or biopsy sample, or *in vitro*, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with *in vitro* model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed  
10 molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number  
15 of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids  
20 in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity (see, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at <http://www.niehs.nih.gov/oc/news/toxchip.htm>). Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

25 In an embodiment, the toxicity of a test compound can be assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with  
30 levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another embodiment relates to the use of the polypeptides disclosed herein to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected

individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of interest. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for MDDT to quantify the levels of MDDT expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) *Anal. Biochem.* 270:103-111; Mendoz, L.G. et al. (1999) *Biotechniques* 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) *Electrophoresis* 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological

sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art (Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662). Various types of microarrays are well known and thoroughly described in Schena, M., ed. (1999; DNA Microarrays: A Practical Approach, Oxford University Press, London).

In another embodiment of the invention, nucleic acid sequences encoding MDDT may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries (Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; Trask, B.J. (1991) Trends Genet. 7:149-154). Once mapped, the nucleic acid sequences may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP) (Lander, E.S. and

D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357).

Fluorescent *in situ* hybridization (FISH) may be correlated with other physical and genetic map data (Heinz-Ulrich, et al. (1995) in Meyers, *supra*, pp. 965-968). Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM)

5 World Wide Web site. Correlation between the location of the gene encoding MDDT on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

*In situ* hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps.

10 Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to  
15 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation (Gatti, R.A. et al. (1988) Nature 336:577-580). The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, MDDT, its catalytic or immunogenic fragments, or  
20 oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between MDDT and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds  
25 having suitable binding affinity to the protein of interest (Geysen, et al. (1984) PCT application WO84/03564). In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with MDDT, or fragments thereof, and washed. Bound MDDT is then detected by methods well known in the art. Purified MDDT can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively,  
30 non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding MDDT specifically compete with a test compound for binding MDDT. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with MDDT.

In additional embodiments, the nucleotide sequences which encode MDDT may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

5       Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

10       The disclosures of all patents, applications, and publications mentioned above and below, including U.S. Ser. No. 60/334,182, U.S. Ser. No. 60/342, 052, U.S. Ser. No. 60/353,284, U.S. Ser. No. 60/350,410, and U.S. Ser. No. 60/363,649, are hereby expressly incorporated by reference.

## EXAMPLES

### I. Construction of cDNA Libraries

15       Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA). Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Invitrogen), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with  
20   chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN,  
25   Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP  
30   vector system (Stratagene) or SUPERScript plasmid system (Invitrogen), using the recommended procedures or similar methods known in the art (Ausubel et al., *supra*, ch. 5). Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000,

SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Biosciences) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Invitrogen, Carlsbad CA), PCDNA2.1 plasmid (Invitrogen), PBK-CMV plasmid (Stratagene), PCR2-  
5 TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pIGEN (Incyte Genomics, Palo Alto CA), pRARE (Incyte Genomics), or pINCY (Incyte Genomics), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 $\alpha$ , DH10B, or ElectroMAX DH10B from Invitrogen.

## II. Isolation of cDNA Clones

10 Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96  
15 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in  
20 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (LabSystems Oy, Helsinki, Finland).

## III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows.  
25 Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Biosciences or supplied in ABI sequencing kits such as the  
30 ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Amersham Biosciences); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading

frames within the cDNA sequences were identified using standard methods (Ausubel et al., *supra*, ch. 7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM; PROTEOME databases with sequences from *Homo sapiens*, *Rattus norvegicus*, *Mus musculus*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Candida albicans* (Incyte Genomics, Palo Alto CA); hidden Markov model (HMM)-based protein family databases such as PFAM, INCY, and TIGRFAM (Haft, D.H. et al. (2001) Nucleic Acids Res. 29:41-43); and HMM-based protein domain databases such as SMART (Schultz, J. et al. (1998) Proc. Natl. Acad. Sci. USA 95:5857-5864; Letunic, I. et al. (2002) Nucleic Acids Res. 30:242-244). (HMM is a probabilistic approach which analyzes consensus primary structures of gene families; see, for example, Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, the PROTEOME databases, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, hidden Markov model (HMM)-based protein family databases such as PFAM, INCY, and TIGRFAM; and HMM-based protein domain databases such as SMART. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (MiraiBio, Alameda CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.



Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:70-138. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 2.

#### **IV. Identification and Editing of Coding Sequences from Genomic DNA**

Putative molecules for disease detection and treatment were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (Burge, C. and S. Karlin (1997) *J. Mol. Biol.* 268:78-94; Burge, C. and S. Karlin (1998) *Curr. Opin. Struct. Biol.* 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode molecules for disease detection and treatment, the encoded polypeptides were analyzed by querying against PFAM models for molecules for disease detection and treatment. Potential molecules for disease detection and treatment were also identified by homology to Incyte cDNA sequences that had been annotated as molecules for disease detection and treatment. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length

polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

## **V. Assembly of Genomic Sequence Data with cDNA Sequence Data**

### **"Stitched" Sequences**

5 Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, 10 generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated 15 but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent 20 type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbprl public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

### **"Stretched" Sequences**

25 Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in 30 Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences

were therefore “stretched” or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

## **VI. Chromosomal Mapping of MDDT Encoding Polynucleotides**

The sequences which were used to assemble SEQ ID NO:70-138 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:70-138 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

### Association of MDDT polynucleotides with Osteoarthritis

Markers that map to regions associated with particular diseases can be used to develop diagnostic and therapeutic tools. Disease association of a chromosome locus is expressed as lod (logarithm of odds) score. The lod score is the logarithm to base 10 of the odds in favor of linkage. Linkage is defined as the tendency of two genes located on the same chromosome to be inherited together through meiosis (*Genetics in Medicine*, Fifth Edition, (1991) Thompson, M.W. et al. W.B. Saunders Co. Philadelphia). A logarithm of the odds ratio for linkage (lod) score of 2 indicates a probability of 1 in 100 that the marker was found solely by chance in affected individuals. In a study of 48 families affected by OA, Loughlin et al. (Rheumatology (2000) 39:377-381) identified D2S202 and D2S117 as two genetic markers with a multiple lod of 2.19 for linkage to OA of the hip.

Restriction fragment length polymorphism (RFLP) markers shown to be near regions of DNA known as sequence-tagged sites (STS), have been mapped to NT\_Contigs generated by the

Human Genome Project using ePCR (Schuler, G.D. (1997) *Genome Research* 7: 541-550, and (1998) *Trends Biotechnol.* 16(11):456-9). Contigs containing regions of DNA with known disease-associated markers are therefore used to identify MDDT sequences that map to disease-associated regions of the genome.

5 Polynucleotides encoding MDDT were mapped to NT\_Contigs. Contigs longer than 1Mb were broken into subcontigs of 1Mb length with overlapping sections of 100kb. A preliminary step used an algorithm, similar to MEGABLAST, to define the mRNA sequence /masked genomic DNA contig pairings. The cDNA/genomic pairings identified by the first algorithm were confirmed, and the MDDT polynucleotides mapped to DNA contigs, using SIM4 (Florea, L. et al. (1998) *Genome*  
10 *Res.* 8:967-74, version May 2000) which had been optimized for high throughput processing and strand assignment confidence). The SIM4 output of the mRNA sequence/genomic contig pairs was further processed to determine the correct location of the MDDT polynucleotides on the genomic contig, as well as their strand identity.

SEQ ID NO:76 was mapped to NT\_029901\_001.3 from Genbank release February 2002,  
15 covering a 6.45 Mb region of the genome that also contains OA-associated genetic markers D2S202 and D2S72. The maximum distance between SEQ ID NO:76 and markers D2S202 and D2S72, therefore, is 6.45 Mb. Thus, SEQ ID NO:76 is in proximity with genetic markers shown to consistently associate with OA. Therefore, in various embodiments, SEQ ID NO:76 can be used for one or more of the following: i) linkage analysis of persons and/or families to the OA disease region  
20 at 2q12-q22, ii) diagnostic assays for osteoarthritis and interleukin expression abnormalities, and iii) developing therapeutics and/or other treatments for OA.

#### Association of MDDT polynucleotides with Lung Cancer

Heritable forms of lung carcinoma have not been reported and thus, identification of relevant disease-associated genes through conventional linkage analysis is not possible.  
25 However, several studies of sporadic nonsmall cell lung carcinoma (NSCLC) tumors have reported loss of heterozygosity (LOH) in regions of chromosome 11 suggesting the presence of one or more tumor suppressor genes (Bepler, G. and Garcia-Blanco, M.A. (1994) *Proc. Natl. Acad. Sci. USA* 91:5513-7; Iizuka, M. (1995) *Genes, Chromosomes & Cancer* 13:40-46; Rasio, D. (1995) *Cancer Research* 55:3988-91). In a study of 79 patients with lung cancer, Iizuka and  
30 coworkers found that 11q14-11q24.2 was deleted in many of the lung tumors studied. Mapping of this region with additional markers showed that the region of chromosome 11q bounded by markers D11S939 and D11S938 was commonly deleted (Iizuka, et al., *supra*). In another study it was shown that human A549 NSCLC cells transformed with a human-derived YAC clone containing a region of chromosome 11q within the region bounded by D11S939 and D11S938,

exhibited little or no increase in cell number (versus control cells whose number increased 5-10-fold in the same 5 day period). Further studies of these hybrid cells showed a decrease in tumorigenicity and an increase in latency following injection into athymic, nude mice, as compared with mice injected with control A549 cells. These studies suggest the presence of a tumor suppressor gene within this region of chromosome 11q and support the association of LOH in this region with NSCLC.

Restriction fragment length polymorphism (RFLP) markers shown to be near regions of DNA known as sequence-tagged sites (STS), have been mapped to NT\_Contigs generated by the Human Genome Project using ePCR (Schuler, G.D. (1997) *Genome Research* 7: 541-550, and (1998) *Trends Biotechnol.* 16(11):456-459). Contigs containing regions of DNA with known disease-associated markers are therefore used to identify MDDT sequences that map to disease-associated regions of the genome.

Polynucleotides encoding MDDT were mapped to NT\_Contigs. Contigs longer than 1Mb were broken into subcontigs of 1Mb length with overlapping sections of 100kb. A preliminary step used an algorithm, similar to MEGABLAST, to define the mRNA sequence /masked genomic DNA contig pairings. The cDNA/genomic pairings identified by the first algorithm were confirmed, and the MDDT polynucleotides mapped to DNA contigs, using SIM4 (Florea, L. et al. (1998) *Genome Res.* 8:967-74, version May 2000) which had been optimized for high throughput processing and strand assignment confidence). The SIM4 output of the mRNA sequence/genomic contig pairs was further processed to determine the correct location of the MDDT polynucleotides on the genomic contig, as well as their strand identity.

SEQ ID NO:102 was mapped to NT\_009151\_019.8 from Genbank release February 2002, covering a 5.5 Mb region of the genome that also contains lung cancer-associated genetic markers D11S939 and D11S938. The maximum distance between SEQ ID NO:102 and markers D11S939 and D11S938, therefore, is 5.5 Mb. Thus, SEQ ID NO:102 is in proximity with genetic markers shown to consistently associate with lung cancer. Therefore, in various embodiments, SEQ ID NO:102 can be used for one or more of the following: i) determination of LOH in persons with lung cancer in the lung cancer disease region at 11q12-24.2, ii) diagnostic assays for lung cancer, and iii) developing therapeutics and/or other treatments for lung cancer.

#### Association of MDDT polynucleotides with Parkinson's Disease

Several genes have been identified as showing linkage to autosomal dominant forms of Parkinson's Disease (PD). PD is a common neurodegenerative disorder causing bradykinesia, resting tremor, muscular rigidity, and postural instability. Cytoplasmic eosinophilic inclusions

called Lewy bodies, and neuronal loss especially in the substantia nigra pars compacta, are pathological hallmarks of PD (Valente, E.M. et al (2001) Am. J. Hum. Genet. 68:895-900). Lewy body Parkinson disease has been thought to be a specific autosomal dominant disorder (Wakabayashi, K. et al. (1998) Acta Neuropath. 96:207-210). Juvenile parkinsonism may be a specific autosomal recessive disorder (Matsumine, H. et al. (1997) Am. J. Hum. Genet. 60: 588-596, 1997). (Online Mendelian Inheritance in Man, OMIM. Johns Hopkins University, Baltimore, MD. MIM Number: 168600: Sept. 9, 2002: . World Wide Web URL: <http://www.ncbi.nlm.nih.gov/omim/>)

Association of a disease with a chromosomal locus can be determined by lod score. Lod score is a statistical method used to test the linkage of two or more loci within families having a genetic disease. The lod score is the logarithm to base 10 of the odds in favor of linkage. Linkage is defined as the tendency of two genes located on the same chromosome to be inherited together through meiosis (*Genetics in Medicine*, Fifth Edition, (1991) Thompson, M.W. et al., W.B. Saunders Co. Philadelphia). A lod score of +3 or greater (1000:1 odds in favor of linkage) indicates a probability of 1 in 1000 that a particular marker was found solely by chance in affected individuals, which is strong evidence that two genetic loci are linked.

One such gene implicated in PD is PARK3, which maps to 2p13 (Gasser, T. et al. (1998) Nature Genet. 18:262-265). A marker at chromosomal position D2S441 was found to have a lod score of 3.2 in the region of PARK3. This marker supported the disease association of PARK3 in the chromosomal interval from D2S134 to D2S286 (Gasser et al., supra). Markers located within chromosomal intervals D2S134 and D2S286, which map between 83.88 to 94.05 centiMorgans on the short arm of chromosome 2, were used to identify genes that map in the region between D2S134 and D2S286.

A second PD gene, implicated in early-onset recessive parkinsonism, is PARK6, located on chromosome 1 at 1p35-1p36. Several markers were obtained with lod scores greater than 3 including D1S199, D1S2732, D1S2828, D1S478, D1S2702, D1S2734, D1S2674 (Valente, E.M. et al, supra). These markers were used to determine the PD-relevant range of chromosome loci and identify sequences that map to chromosome 1 between D1S199 and D1S2885.

Restriction fragment length polymorphism (RFLP) markers shown to be near regions of DNA known as sequence-tagged sites (STS), have been mapped to NT\_Contigs generated by the Human Genome Project using ePCR (Schuler, G.D. (1997) Genome Research 7: 541-550, and (1998) Trends Biotechnol. 16(11):456-9). Contigs containing regions of DNA with known

disease-associated markers are therefore used to identify MDDT sequences that map to disease-associated regions of the genome.

Polynucleotides encoding MDDT were mapped to NT\_Contigs. Contigs longer than 1Mb were broken into subcontigs of 1Mb length with overlapping sections of 100kb. A preliminary step used an algorithm, similar to MEGABLAST, to define the mRNA sequence /masked genomic DNA contig pairings. The cDNA/genomic pairings identified by the first algorithm were confirmed, and the MDDT polynucleotides mapped to DNA contigs, using SIM4 (Florea, L. et al. (1998) Genome Res. 8:967-74, version May 2000) which had been optimized for high throughput processing and strand assignment confidence). The SIM4 output of the mRNA sequence/genomic contig pairs was further processed to determine the correct location of the MDDT polynucleotides on the genomic contig, as well as their strand identity.

SEQ ID NO:131 was mapped to NT\_025651\_003.7 from Genbank release February 2002, covering a 9.65 Mb region of the genome that also contains PD-associated genetic markers D2S134 and D2S286. The maximum distance between SEQ ID NO:131 and markers D2S134 and D2S286, therefore, is 9.65 Mb. Thus, SEQ ID NO:131 is in proximity with genetic markers shown to consistently associate with PD.

In an alternative example, SEQ ID NO:134 was mapped to NT\_025651\_002.7 from Genbank release February 2002, covering a 9.65 Mb region of the genome that also contains PD-associated genetic markers D2S134 and D2S286. The maximum distance between SEQ ID NO:134 and markers D2S134 and D2S286, therefore, is 9.65 Mb. Thus, SEQ ID NO:134 is in proximity with genetic markers shown to consistently associate with PD.

## VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound (Sambrook and Russell, *supra*, ch. 7; Ausubel et al., *supra*, ch. 4).

Analogous computer techniques applying BLAST were used to search for identical or related molecules in databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum} \{ \text{length}(\text{Seq. 1}), \text{length}(\text{Seq. 2}) \}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotides encoding MDDT are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding MDDT. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

#### **VIII. Extension of MDDT Encoding Polynucleotides**

Full length polynucleotides are produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was



synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at  
5 temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR  
10 was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing  $Mg^{2+}$ ,  $(NH_4)_2SO_4$ , and 2-mercaptoethanol, Taq DNA polymerase (Amersham Biosciences), ELONGASE enzyme (Invitrogen), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2  
15 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100  $\mu$ l PICOGREEN  
20 quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5  $\mu$ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5  $\mu$ l to 10  $\mu$ l aliquot of the reaction mixture was analyzed by  
25 electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Biosciences). For shotgun  
30 sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Biosciences), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing

media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Biosciences) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Biosciences) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotides are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

#### 15 **IX. Identification of Single Nucleotide Polymorphisms in MDDT Encoding Polynucleotides**

Common DNA sequence variants known as single nucleotide polymorphisms (SNPs) were identified in SEQ ID NO:70-138 using the LIFESEQ database (Incyte Genomics). Sequences from the same gene were clustered together and assembled as described in Example III, allowing the identification of all sequence variants in the gene. An algorithm consisting of a series of filters was used to distinguish SNPs from other sequence variants. Preliminary filters removed the majority of basecall errors by requiring a minimum Phred quality score of 15, and removed sequence alignment errors and errors resulting from improper trimming of vector sequences, chimeras, and splice variants. An automated procedure of advanced chromosome analysis analysed the original chromatogram files in the vicinity of the putative SNP. Clone error filters used statistically generated algorithms to identify errors introduced during laboratory processing, such as those caused by reverse transcriptase, polymerase, or somatic mutation. Clustering error filters used statistically generated algorithms to identify errors resulting from clustering of close homologs or pseudogenes, or due to contamination by non-human sequences. A final set of filters removed duplicates and SNPs found in immunoglobulins or T-cell receptors.

30 Certain SNPs were selected for further characterization by mass spectrometry using the high throughput MASSARRAY system (Sequenom, Inc.) to analyze allele frequencies at the SNP sites in four different human populations. The Caucasian population comprised 92 individuals (46 male, 46 female), including 83 from Utah, four French, three Venezuelan, and two Amish individuals. The African population comprised 194 individuals (97 male, 97 female), all African Americans. The

Hispanic population comprised 324 individuals (162 male, 162 female), all Mexican Hispanic. The Asian population comprised 126 individuals (64 male, 62 female) with a reported parental breakdown of 43% Chinese, 31% Japanese, 13% Korean, 5% Vietnamese, and 8% other Asian. Allele frequencies were first analyzed in the Caucasian population; in some cases those SNPs which showed  
5 no allelic variance in this population were not further tested in the other three populations.

#### **X. Labeling and Use of Individual Hybridization Probes**

Hybridization probes derived from SEQ ID NO:70-138 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide  
10 fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250  $\mu$ Ci of [ $\gamma$ - $^{32}$ P] adenosine triphosphate (Amersham Biosciences), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25  
superfine size exclusion dextran bead column (Amersham Biosciences). An aliquot containing 10<sup>7</sup>  
15 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16  
20 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

#### **XI. Microarrays**

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing; see, e.g., Baldeschweiler et al., *supra*), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the  
25 aforementioned technologies should be uniform and solid with a non-porous surface (Skena, M., ed. (1999) DNA Microarrays: A Practical Approach, Oxford University Press, London). Suggested  
30 substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements (Skena, M. et al. (1995) *Science* 270:467-470;

Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31).

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

#### **Tissue or Cell Sample Preparation**

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)<sup>+</sup> RNA is purified using the oligo-(dT) cellulose method. Each poly(A)<sup>+</sup> RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ $\mu$ l oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/ $\mu$ l RNase inhibitor, 500  $\mu$ M dATP, 500  $\mu$ M dGTP, 500  $\mu$ M dTTP, 40  $\mu$ M dCTP, 40  $\mu$ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Biosciences). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)<sup>+</sup> RNA with GEMBRIGHT kits (Incyte Genomics). Specific control poly(A)<sup>+</sup> RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (Clontech, Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14  $\mu$ l 5X SSC/0.2% SDS.

#### **Microarray Preparation**

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are

amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5  $\mu$ g. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Biosciences).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in U.S. Patent No. 5,807,522, incorporated herein by reference. 1  $\mu$ l of the array element DNA, at an average concentration of 100 ng/ $\mu$ l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60°C followed by washes in 0.2% SDS and distilled water as before.

#### **Hybridization**

Hybridization reactions contain 9  $\mu$ l of sample mixture consisting of 0.2  $\mu$ g each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm<sup>2</sup> coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140  $\mu$ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

#### **Detection**

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-

scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte Genomics). Array elements that exhibit at least about a two-fold change in expression, a signal-to-background ratio of at least about 2.5, and an element spot size of at least about 40%, are considered to be differentially expressed.

#### Expression

For example, SEQ ID NO:77 and SEQ ID NO:81 showed differential expression, as determined by microarray analysis, in liver C3A cells treated with one of the following steroids: beclomethasone, budesonide, dexamethasone, and betamethasone. The human C3A cell line is a clonal derivative of HepG2/C3 (hepatoma cell line, isolated from a 15-year-old male with liver tumor), which was selected for strong contact inhibition of growth. The use of a clonal population enhances the reproducibility of the cells. C3A cells have many characteristics of primary human hepatocytes in culture: i) expression of insulin receptor and insulin-like growth factor II receptor; ii) secretion of a high ratio of serum albumin compared with  $\alpha$ -fetoprotein iii) conversion of ammonia to urea and glutamine; iv) metabolism of aromatic amino acids; and v) proliferation in glucose-free and insulin-free medium. The C3A cell line is now well established as an *in vitro* model of the mature human liver (Mickelson et al. (1995) Hepatology 22:866-875; Nagendra et al. (1997) Am J Physiol 272:G408-G416). C3A cells were treated with 1, 10, and 100  $\mu$ M beclomethasone, budesonide, dexamethasone, and betamethasone for 1hr, 3hr, 6hr and were compared with untreated cells. In one experiment, SEQ ID NO:77 and 81 showed at least a two-fold increase in expression at a minimum of two out of the three time points in C3A cells treated with beclomethasone. In a separate experiment, SEQ ID NO:77 and 81 showed at least a two-fold increase in expression in C3A cells treated with 1  $\mu$ M budesonide for 1, 3, and 6 hours. SEQ ID NO:77 and SEQ ID NO:81 also showed at least a two-fold increase in expression in C3A cells treated with 100  $\mu$ M dexamethasone and betamethasone for 1, 3, and 6 hours. These experiments indicate that SEQ ID NO:77 and SEQ ID NO:81 are useful for monitoring the pharmacodynamics of drugs and effects on liver metabolism upon steroid therapy.

As another example, SEQ ID NO:79 showed differential expression in MCF7 breast carcinoma cell line versus primary mammary epithelial cells as determined by microarray analysis. In three separate experiments, the gene expression profile of a primary mammary epithelial cell line, HMEC, was compared to the gene expression profiles of breast carcinoma lines at different stages of tumor progression. Cell lines compared included: a) MCF-10A, a breast mammary gland cell line isolated from a 36-year-old woman with fibrocystic breast disease; b) MCF7, a nonmalignant breast adenocarcinoma cell line isolated from the pleural effusion of a 69-year-old female; c) T-47D, a breast carcinoma cell line isolated from a pleural effusion obtained from a 54-year-old female with an infiltrating ductal carcinoma of the breast; d) Sk-BR-3, a breast adenocarcinoma cell line isolated from a malignant pleural effusion of a 43-year-old female; e) BT-20, a breast carcinoma cell line derived *in vitro* from tumor mass isolated from a 74-year-old female; f) MDA-mb-231, a breast tumor cell line isolated from the pleural effusion of a 51-year old female; g) MDA-mb-435S, a spindle shaped strain that evolved from the parent line (435) isolated from the pleural effusion of a 31-year-old female with metastatic, ductal adenocarcinoma of the breast; h) BT20, a breast carcinoma cell line

derived in vitro from cells emigrating out of thin slices of a tumor mass isolated from a 74-year-old female; i) BT474, a breast ductal carcinoma cell line isolated from a solid, invasive ductal carcinoma of the breast from a 60-year-old female; j) BT483, a breast ductal carcinoma cell line isolated from a papillary invasive ductal tumor from a 23-year-old normal, menstruating, parous female; k) HS578T, a breast ductal carcinoma cell line isolated from a 74-year-old female with breast carcinoma;; and l) MDA-mb-468, a breast adenocarcinoma cell line isolated from the pleural effusion of a 51-year-old female with metastatic adenocarcinoma of the breast. The microarray experiments showed that in all three experiments, the expression of SEQ ID NO:79 was increased by at least two fold in cells from the MCF7 cell line relative to cells from the primary mammary epithelial cell line, HMEC.

Therefore, SEQ ID NO:79 is useful in diagnostic assays for and monitoring treatment of breast cancer.

In an alternative example, gene expression profiles of nonmalignant mammary epithelial cells were compared to gene expression profiles of various breast carcinoma lines at different stages of tumor progression. The cells were grown in defined serum-free H14 medium to 70-80% confluence prior to RNA harvest. Cell lines compared included: a) HMEC, a primary breast epithelial cell line isolated from a normal donor, b) MCF-10A, a breast mammary gland cell line isolated from a 36-year-old woman with fibrocystic breast disease, c) MCF7, a nonmalignant breast adenocarcinoma cell line isolated from the pleural effusion of a 69-year-old female, d) T-47D, a breast carcinoma cell line isolated from a pleural effusion obtained from a 54-year-old female with an infiltrating ductal carcinoma of the breast, e) Sk-BR-3, a breast adenocarcinoma cell line isolated from a malignant pleural effusion of a 43-year-old female, f) BT20, a breast carcinoma cell line derived *in vitro* from cells emigrating out of thin slices of the tumor mass isolated from a 74-year-old female, g) MDA-mb-231, a breast tumor cell line isolated from the pleural effusion of a 51-year-old female, and h) MDA-mb-435S, a spindle-shaped strain that evolved from the parent line (435) isolated by R. Cailleau from pleural effusion of a 31-year-old female with metastatic, ductal adenocarcinoma of the breast.

The microarray experiments described above showed that expression of both SEQ ID NO:83 and SEQ ID NO:93 were decreased by at least two fold in the MDA-mb-435S cell line relative to primary mammary epithelial cells. These experiments also demonstrated that expression of SEQ ID NO:86 was decreased by at least two fold in the BT20 and MCF7 cell lines relative to primary mammary epithelial cells. Therefore, SEQ ID NO:83, SEQ ID NO:86, and SEQ ID NO:93 are useful in diagnostic and staging assays for breast cancer and as potential biological markers and therapeutic agents in the treatment of breast cancer.

In an alternative example, gene expression profiles were obtained by comparing normal colon tissue from a 56-year-old female diagnosed with poorly differentiated metastatic adenocarcinoma of



possible ovarian origin and a clinical history of recurrent cecal mass, to associated colon tumor tissue from the same donor (Huntsman Cancer Institute, Salt Lake City, UT) by competitive hybridization.

These experiments showed that expression of SEQ ID NO:83 and SEQ ID NO:93 were both increased by more than two fold in colon adenocarcinoma tissue as compared to normal colon tissue.

5 Therefore, SEQ ID NO:83 and SEQ ID NO:93 are useful in diagnostic assays for colon cancer and as a potential biological marker and therapeutic agent in the treatment of colon cancer.

In an alternative example, SEQ ID NO:94 showed differential expression, as determined by microarray analysis, in Alzheimer's Disease (AD). In a comparison of posterior cingulate tissue from a 68-year-old female with mild AD to posterior cingulate tissue from a normal 61-year-old female, the  
10 expression of SEQ ID NO:94 was increased at least two-fold. Therefore, SEQ ID NO:94 is useful in diagnostic assays for AD and as a potential biological marker and therapeutic agent in the treatment of AD.

In an alternative example, SEQ ID NO:115 and SEQ ID NO:125 were differentially expressed in human colon tumor tissue as compared to normal colon tissue from the same donors.  
15 Colon cancer evolves through a multi-step process whereby pre-malignant colonocytes undergo a relatively defined sequence of events leading to tumor formation. Several factors participate in the process of tumor progression and malignant transformation including genetic factors, mutations, and selection. Despite efforts to characterize the molecular events leading to colon cancer, the process of tumor development and progression has not been delineated. To identify genes differentially  
20 expressed in colon cancer, gene expression patterns in normal and tumor tissues from the same donor were compared using competitive hybridization. This process eliminates some of the individual variation due to genetic background, and enhances differences due to the disease process.

SEQ ID NO:115 and SEQ ID NO:125 were underexpressed by at least two-fold in the colon tumor tissue. These experiments indicate that SEQ ID NO:115 and SEQ ID NO:125 exhibited  
25 significant differential expression patterns using microarray techniques, and further establish their utility as diagnostic markers or therapeutic agents which may be useful in a variety of conditions and diseases, including colon cancer.

## **XII. Complementary Polynucleotides**

Sequences complementary to the MDDT-encoding sequences, or any parts thereof, are used  
30 to detect, decrease, or inhibit expression of naturally occurring MDDT. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of MDDT. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence

and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the MDDT-encoding transcript.

### **XIII. Expression of MDDT**

5           Expression and purification of MDDT is achieved using bacterial or virus-based expression systems. For expression of MDDT in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid  
10   element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express MDDT upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of MDDT in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant *Autographica californica* nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is  
15   replaced with cDNA encoding MDDT by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect *Spodoptera frugiperda* (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus (Engelhard, E.K. et al.  
20   (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945).

          In most expression systems, MDDT is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-  
25   kilodalton enzyme from *Schistosoma japonicum*, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Biosciences). Following purification, the GST moiety can be proteolytically cleaved from MDDT at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-  
30   His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel et al. (*supra*, ch. 10 and 16). Purified MDDT obtained by these methods can be used directly in the assays shown in Examples XVII, XVIII, and XIX, where applicable.

### **XIV. Functional Assays**

MDDT function is assessed by expressing the sequences encoding MDDT at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT plasmid (Invitrogen, Carlsbad CA) and PCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10  $\mu$ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2  $\mu$ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994; Flow Cytometry, Oxford, New York NY).

The influence of MDDT on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding MDDT and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding MDDT and other genes of interest can be analyzed by northern analysis or microarray techniques.

#### **XV. Production of MDDT Specific Antibodies**

MDDT substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize animals (e.g., rabbits, mice, etc.) and to produce antibodies using standard protocols.

Alternatively, the MDDT amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art (Ausubel et al., *supra*, ch. 11).

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using Fmoc chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity (Ausubel et al., *supra*). Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for anti-peptide and anti-MDDT activity by, for example, binding the peptide or MDDT to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

#### **XVI. Purification of Naturally Occurring MDDT Using Specific Antibodies**

Naturally occurring or recombinant MDDT is substantially purified by immunoaffinity chromatography using antibodies specific for MDDT. An immunoaffinity column is constructed by covalently coupling anti-MDDT antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Biosciences). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing MDDT are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of MDDT (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/MDDT binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and MDDT is collected.

#### **XVII. Identification of Molecules Which Interact with MDDT**

MDDT, or biologically active fragments thereof, are labeled with <sup>125</sup>I Bolton-Hunter reagent (Bolton, A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539). Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled MDDT, washed, and any wells with labeled MDDT complex are assayed. Data obtained using different concentrations of MDDT are used to calculate values for the number, affinity, and association of MDDT with the candidate molecules.

Alternatively, molecules interacting with MDDT are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989; *Nature* 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

MDDT may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

## 5 XVIII. Demonstration of MDDT Activity

Phorbol ester binding activity of MDDT is measured using an assay based on the fluorescent phorbol ester sapinotoxin-D (SAPD). Binding of SAPD to MDDT is quantified by measuring the resonance energy transfer from MDDT tryptophans to the 2-(N-methylamino)benzoyl fluorophore of the phorbol ester, as described by Slater et al. ((1996) J. Biol. Chem. 271:4627-4631).

10 Another assay for MDDT activity measures the expression of MDDT on the cell surface. cDNA encoding MDDT is transfected into an appropriate mammalian cell line. Cell surface proteins are labeled with biotin as described (de la Fuente, M.A. et al. (1997) Blood 90:2398-2405). Immunoprecipitations are performed using MDDT-specific antibodies, and immunoprecipitated samples are analyzed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of MDDT expressed on the cell surface.

In the alternative, an assay for MDDT activity is based on a prototypical assay for ligand/receptor-mediated modulation of cell proliferation. This assay measures the rate of DNA synthesis in Swiss mouse 3T3 cells. A plasmid containing polynucleotides encoding MDDT is added to quiescent 3T3 cultured cells using transfection methods well known in the art. The transiently transfected cells are then incubated in the presence of [<sup>3</sup>H]thymidine, a radioactive DNA precursor molecule. Varying amounts of MDDT ligand are then added to the cultured cells. Incorporation of [<sup>3</sup>H]thymidine into acid-precipitable DNA is measured over an appropriate time interval using a radioisotope counter, and the amount incorporated is directly proportional to the amount of newly synthesized DNA. A linear dose-response curve over at least a hundred-fold MDDT ligand concentration range is indicative of receptor activity. One unit of activity per milliliter is defined as the concentration of MDDT producing a 50% response level, where 100% represents maximal incorporation of [<sup>3</sup>H]thymidine into acid-precipitable DNA (McKay, I. and I. Leigh, eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York NY, p. 73.)

30 In a further alternative, the assay for MDDT activity is based upon the ability of GPCR family proteins to modulate G protein-activated second messenger signal transduction pathways (e.g., cAMP; Gaudin, P. et al. (1998) J. Biol. Chem. 273:4990-4996). A plasmid encoding full length MDDT is transfected into a mammalian cell line (e.g., Chinese hamster ovary (CHO) or human embryonic kidney (HEK-293) cell lines) using methods well-known in the art. Transfected cells are

grown in 12-well trays in culture medium for 48 hours, then the culture medium is discarded, and the attached cells are gently washed with PBS. The cells are then incubated in culture medium with or without ligand for 30 minutes, then the medium is removed and cells lysed by treatment with 1 M perchloric acid. The cAMP levels in the lysate are measured by radioimmunoassay using methods well-known in the art. Changes in the levels of cAMP in the lysate from cells exposed to ligand compared to those without ligand are proportional to the amount of MDDT present in the transfected cells.

To measure changes in inositol phosphate levels, the cells are grown in 24-well plates containing  $1 \times 10^5$  cells/well and incubated with inositol-free media and [ $^3\text{H}$ ]myoinositol, 2  $\mu\text{Ci}$ /well, for 48 hr. The culture medium is removed, and the cells washed with buffer containing 10 mM LiCl followed by addition of ligand. The reaction is stopped by addition of perchloric acid. Inositol phosphates are extracted and separated on Dowex AG1-X8 (Bio-Rad) anion exchange resin, and the total labeled inositol phosphates counted by liquid scintillation. Changes in the levels of labeled inositol phosphate from cells exposed to ligand compared to those without ligand are proportional to the amount of MDDT present in the transfected cells.

In a further alternative, the ion conductance capacity of MDDT is demonstrated using an electrophysiological assay. MDDT is expressed by transforming a mammalian cell line such as COS7, HeLa or CHO with a eukaryotic expression vector encoding MDDT. Eukaryotic expression vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. A small amount of a second plasmid, which expresses any one of a number of marker genes such as  $\beta$ -galactosidase, is co-transformed into the cells in order to allow rapid identification of those cells which have taken up and expressed the foreign DNA. The cells are incubated for 48-72 hours after transformation under conditions appropriate for the cell line to allow expression and accumulation of MDDT and  $\beta$ -galactosidase. Transformed cells expressing  $\beta$ -galactosidase are stained blue when a suitable colorimetric substrate is added to the culture media under conditions that are well known in the art. Stained cells are tested for differences in membrane conductance due to various ions by electrophysiological techniques that are well known in the art. Untransformed cells, and/or cells transformed with either vector sequences alone or  $\beta$ -galactosidase sequences alone, are used as controls and tested in parallel. The contribution of MDDT to cation or anion conductance can be shown by incubating the cells using antibodies specific for either MDDT. The respective antibodies will bind to the extracellular side of MDDT, thereby blocking the pore in the ion channel, and the associated conductance.

In a further alternative, MDDT transport activity is assayed by measuring uptake of labeled substrates into Xenopus laevis oocytes. Oocytes at stages V and VI are injected with MDDT mRNA

(10 ng per oocyte) and incubated for 3 days at 18 °C in OR2 medium (82.5 mM NaCl, 2.5 mM KCl, 1 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, 1 mM Na<sub>2</sub>HPO<sub>4</sub>, 5 mM Hepes, 3.8 mM NaOH, 50 µg/ml gentamycin, pH 7.8) to allow expression of MDDT protein. Oocytes are then transferred to standard uptake medium (100 mM NaCl, 2 mM KCl, 1 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, 10 mM Hepes/Tris pH 7.5). Uptake of various substrates (e.g., amino acids, sugars, drugs, and neurotransmitters) is initiated by adding a <sup>3</sup>H substrate to the oocytes. After incubating for 30 minutes, uptake is terminated by washing the oocytes three times in Na<sup>+</sup>-free medium, measuring the incorporated <sup>3</sup>H, and comparing with controls. MDDT activity is proportional to the level of internalized <sup>3</sup>H substrate.

In a further alternative, MDDT protein kinase (PK) activity is measured by phosphorylation of a protein substrate using gamma-labeled [<sup>32</sup>P]-ATP and quantitation of the incorporated radioactivity using a gamma radioisotope counter. MDDT is incubated with the protein substrate, [<sup>32</sup>P]-ATP, and an appropriate kinase buffer. The <sup>32</sup>P incorporated into the product is separated from free [<sup>32</sup>P]-ATP by electrophoresis and the incorporated <sup>32</sup>P is counted. The amount of <sup>32</sup>P recovered is proportional to the PK activity of MDDT in the assay. A determination of the specific amino acid residue phosphorylated is made by phosphoamino acid analysis of the hydrolyzed protein.

In a further alternative, an *in vitro* assay for MDDT activity measures the transformation of normal human fibroblast cells overexpressing antisense MDDT RNA. (Garkavtsev, I. and Riabowol, K. (1997) Mol. Cell Biol. 17:2014-2019.) cDNA encoding MDDT is subcloned into the pLNCX retroviral vector to enable expression of antisense MDDT RNA. The resulting construct is transfected into the ecotropic BOSC23 virus-packaging cell line. Virus contained in the BOSC23 culture supernatant is used to infect the amphotropic CAK8 virus-packaging cell line. Virus contained in the CAK8 culture supernatant is used to infect normal human fibroblast (Hs68) cells. Infected cells are assessed for the following quantifiable properties characteristic of transformed cells: growth in culture to high density associated with loss of contact inhibition, growth in suspension or in soft agar, formation of colonies or foci, lowered serum requirements, and ability to induce tumors when injected into immunodeficient mice. The activity of MDDT is proportional to the extent of transformation of Hs68 cells.

Alternatively, MDDT can be expressed in a mammalian cell line by transforming the cells with a eukaryotic expression vector encoding MDDT. Eukaryotic expression vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. To assay the cellular localization of MDDT, cells are fractionated as described by Jiang H. P. et al. (1992; Proc. Natl. Acad. Sci. 89: 7856-7860). Briefly, cells pelleted by low-speed centrifugation are resuspended in buffer (10 mM TRIS-HCl, pH 7.4/ 10mM NaCl/ 3mM MgCl<sub>2</sub>/ 5mM EDTA with 10ug/ml aprotinin, 10ug/ml leupeptin, 10ug/ml pepstatin A, 0.2mM

phenylmethylsulfonyl fluoride) and homogenized. The homogenate is centrifuged at 600 x g for 5 minutes. The particulate and cytosol fractions are separated by ultracentrifugation of the supernatant at 100,000 x g for 60 minutes. The nuclear fraction is obtained by resuspending the 600 x g pellet in sucrose solution (0.25 M sucrose/ 10mM TRIS-HCl, pH 7.4/ 2mM MgCl<sub>2</sub>) and recentrifuged at 600 x g. Equal amounts of protein from each fraction are applied to an SDS/10% polyacrylamide gel and blotted onto membranes. Western blot analysis is performed using MDDT anti-serum. The localization of MDDT is assessed by the intensity of the corresponding band in the nuclear fraction relative to the intensity in the other fractions. Alternatively, the presence of MDDT in cellular fractions is examined by fluorescence microscopy using a fluorescent antibody specific for MDDT.

Alternatively, MDDT activity may be demonstrated as the ability to interact with its associated Ras superfamily protein, in an in vitro binding assay. The candidate Ras superfamily proteins are expressed as fusion proteins with glutathione S-transferase (GST), and purified by affinity chromatography on glutathione-Sepharose. The Ras superfamily proteins are loaded with GDP by incubating 20 mM Tris buffer, pH 8.0, containing 100 mM NaCl, 2 mM EDTA, 5 mM MgCl<sub>2</sub>, 0.2 mM DTT, 100 μM AMP-PNP and 10 μM GDP at 30°C for 20 minutes. MDDT is expressed as a FLAG fusion protein in a baculovirus system. Extracts of these baculovirus cells containing MDDT-FLAG fusion proteins are precleared with GST beads, then incubated with GST-Ras superfamily fusion proteins. The complexes formed are precipitated by glutathione-Sepharose and separated by SDS-polyacrylamide gel electrophoresis. The separated proteins are blotted onto nitrocellulose membranes and probed with commercially available anti-FLAG antibodies. MDDT activity is proportional to the amount of MDDT-FLAG fusion protein detected in the complex.

Alternatively, MDDT activity is demonstrated by measuring the induction of terminal differentiation or cell cycle progression when MDDT is expressed at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT™ (Life Technologies, Gaithersburg, MD) and pCR™ 3.1 (Invitrogen, Carlsbad, CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP) (Clontech, Palo Alto, CA), CD64, or a CD64-GFP fusion protein. Flow cytometry detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell



cycle progression or terminal differentiation. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; up or down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M. G. (1994) Flow Cytometry, Oxford, New York, NY.

#### **XIX. Identification of MDDT Ligands**

MDDT is expressed in a eukaryotic cell line such as CHO (Chinese Hamster Ovary) or HEK (Human Embryonic Kidney) 293 which have a good history of GPCR expression and which contain a wide range of G-proteins allowing for functional coupling of the expressed MDDT to downstream effectors. The transformed cells are assayed for activation of the expressed receptors in the presence of candidate ligands. Activity is measured by changes in intracellular second messengers, such as cyclic AMP or  $\text{Ca}^{2+}$ . These may be measured directly using standard methods well known in the art, or by the use of reporter gene assays in which a luminescent protein (e.g. firefly luciferase or green fluorescent protein) is under the transcriptional control of a promoter responsive to the stimulation of protein kinase C by the activated receptor (Milligan, G. et al. (1996) Trends Pharmacol. Sci. 17:235-237). Assay technologies are available for both of these second messenger systems to allow high throughput readout in multi-well plate format, such as the adenylyl cyclase activation FlashPlate Assay (NEN Life Sciences Products), or fluorescent  $\text{Ca}^{2+}$  indicators such as Fluo-4 AM (Molecular Probes) in combination with the FLIPR fluorimetric plate reading system (Molecular Devices). In cases where the physiologically relevant second messenger pathway is not known, MDDT may be coexpressed with the G-proteins  $\text{G}_{\alpha 15/16}$  which have been demonstrated to couple to a wide range of G-proteins (Offermanns, S. and M.I. Simon (1995) J. Biol. Chem. 270:15175-15180), in order to funnel the signal transduction of the MDDT through a pathway involving phospholipase C and  $\text{Ca}^{2+}$  mobilization. Alternatively, MDDT may be expressed in engineered yeast systems which lack endogenous GPCRs, thus providing the advantage of a null background for MDDT activation screening. These yeast systems substitute a human GPCR and  $\text{G}_{\alpha}$  protein for the corresponding components of the endogenous yeast pheromone receptor pathway. Downstream signaling pathways are also modified so that the normal yeast response to the signal is converted to positive growth on selective media or to reporter gene expression (Broach, J.R. and J. Thorner (1996) Nature 384 (supp.):14-16). The receptors are screened against putative ligands including known GPCR ligands

and other naturally occurring bioactive molecules. Biological extracts from tissues, biological fluids and cell supernatants are also screened.

Various modifications and variations of the described compositions, methods, and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. It will be appreciated that the invention provides novel and useful proteins, and their encoding polynucleotides, which can be used in the drug discovery process, as well as methods for using these compositions for the detection, diagnosis, and treatment of diseases and conditions. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Nor should the description of such embodiments be considered exhaustive or limit the invention to the precise forms disclosed. Furthermore, elements from one embodiment can be readily recombined with elements from one or more other embodiments. Such combinations can form a number of embodiments within the scope of the invention. It is intended that the scope of the invention be defined by the following claims and their equivalents.

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
1393336	1	1393336CD1	70	1393336CB1
1431502	2	1431502CD1	71	1431502CB1
2445220	3	2445220CD1	72	2445220CB1
5504385	4	5504385CD1	73	5504385CB1
6974948	5	6974948CD1	74	6974948CB1
7501636	6	7501636CD1	75	7501636CB1
2535717	7	2535717CD1	76	2535717CB1
6119548	8	6119548CD1	77	6119548CB1
72263451	9	72263451CD1	78	72263451CB1
7502640	10	7502640CD1	79	7502640CB1
7505807	11	7505807CD1	80	7505807CB1
7506413	12	7506413CD1	81	7506413CB1
1283631	13	1283631CD1	82	1283631CB1
1740413	14	1740413CD1	83	1740413CB1
1951731	15	1951731CD1	84	1951731CB1
3741930	16	3741930CD1	85	3741930CB1
5402506	17	5402506CD1	86	5402506CB1
71081333	18	71081333CD1	87	71081333CB1
7503139	19	7503139CD1	88	7503139CB1
7505836	20	7505836CD1	89	7505836CB1
7505858	21	7505858CD1	90	7505858CB1
7505872	22	7505872CD1	91	7505872CB1
7506456	23	7506456CD1	92	7506456CB1
7506697	24	7506697CD1	93	7506697CB1
7623472	25	7623472CD1	94	7623472CB1
7506416	26	7506416CD1	95	7506416CB1
4823849	27	4823849CD1	96	4823849CB1
4433922	28	4433922CD1	97	4433922CB1
7504597	29	7504597CD1	98	7504597CB1
7505987	30	7505987CD1	99	7505987CB1
7506025	31	7506025CD1	100	7506025CB1
7506102	32	7506102CD1	101	7506102CB1
1333949	33	1333949CD1	102	1333949CB1
7035533	34	7035533CD1	103	7035533CB1
2815375	35	2815375CD1	104	2815375CB1
2820152	36	2820152CD1	105	2820152CB1
2959305	37	2959305CD1	106	2959305CB1
4913449	38	4913449CD1	107	4913449CB1
7506136	39	7506136CD1	108	7506136CB1
7506225	40	7506225CD1	109	7506225CB1
7506227	41	7506227CD1	110	7506227CB1
3144431	42	3144431CD1	111	3144431CB1
2633315	43	2633315CD1	112	2633315CB1
3401751	44	3401751CD1	113	3401751CB1
045680	45	045680CD1	114	045680CB1
1503172	46	1503172CD1	115	1503172CB1
1818665	47	1818665CD1	116	1818665CB1
3251352	48	3251352CD1	117	3251352CB1

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
55091643	49	55091643CD1	118	55091643CB1
7500770	50	7500770CD1	119	7500770CB1
7506350	51	7506350CD1	120	7506350CB1
7508370	52	7508370CD1	121	7508370CB1
2894093	53	2894093CD1	122	2894093CB1
7507335	54	7507335CD1	123	7507335CB1
7509081	55	7509081CD1	124	7509081CB1
7502450	56	7502450CD1	125	7502450CB1
7501405	57	7501405CD1	126	7501405CB1
7504528	58	7504528CD1	127	7504528CB1
7509049	59	7509049CD1	128	7509049CB1
7509086	60	7509086CD1	129	7509086CB1
7506914	61	7506914CD1	130	7506914CB1
5606114	62	5606114CD1	131	5606114CB1
7503282	63	7503282CD1	132	7503282CB1
7503284	64	7503284CD1	133	7503284CB1
7510501	65	7510501CD1	134	7510501CB1
7500444	66	7500444CD1	135	7500444CB1
7510297	67	7510297CD1	136	7510297CB1
7640560	68	7640560CD1	137	7640560CB1
7506087	69	7506087CD1	138	7506087CB1

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
1	1393336CD1	g14150747 700874 LOC81558	2.5E-50 8.4E-10	[Mus musculus] GIG18 [Homo sapiens] C/EBP-induced protein
2	1431502CD1	341072 PKP2	1.6E-11	[Homo sapiens][Nuclear; Plasma membrane; Cell junction] Plakophilin 2, a member of the armadillo family of junctional plaque proteins, localizes to both desmosomes and to nuclear particles that contain RNA Polymerase III and the transcription factor TFIIB (Mertens, C. et al. (1996) Plakophilins 2a and 2b: constitutive proteins of dual location in the karyoplasm and the desmosomal plaque. J. Cell Biol. 135:1009-1025; Mertens, C. et al. (2001) Nuclear particles containing RNA polymerase III complexes associated with the junctional plaque protein plakophilin 2. Proc. Natl. Acad. Sci. U S A 98:7795-7800.)
5	6974948CD1	370403 SPBC20 F10.07	2.2E-32	[Schizosaccharomyces pombe] Protein which appears to localize to membranes (Ding, D. Q. et al. (2000) Large-scale screening of intracellular protein localization in living fission yeast cells by the use of a GFP-fusion genomic DNA library. Genes Cells 5:169-190).
7	2535717CD1	g15823659	1.6E-129	[Homo sapiens] (AB053317) ALS2CR15 (Hadano, S. et al. (2001) A gene encoding a putative GTPase regulator is mutated in familial amyotrophic lateral sclerosis 2. Nat. Genet. 29:166-173.)

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		661206 ICA1	2.8E-102	[Homo sapiens][Cytoplasmic] Islet cell autoantigen, an autoantigen in type 1 insulin dependent diabetes (Pugliese, A. et al. (1997) The insulin gene is transcribed in the human thymus and transcription levels correlated with allelic variation at the INS VNTR-IDDM2 susceptibility locus for type 1 diabetes. Nat Genet 15:293-297).
8	6119548CD1	g15192149	1.0E-172	[Rattus norvegicus] IIIG9 long form
9	72263451CD1	g14550467	7.9E-11	[Homo sapiens] spermatogenesis associated 2
		343030 SPATA2	6.9E-12	[Homo sapiens][Cytoplasmic] Spermatogenesis associated 2, a cytoplasmic protein of Sertoli cells which may have a role in spermatogenesis; expression is altered in various testiculopathies (Graziotto, R. et al. (1999) cDNA cloning and characterization of PD1: a novel human testicular protein with different expressions in various testiculopathies. Exp. Cell Res. 248:620-626.)
10	7502640CD1	g11345415	6.0E-207	[Mus musculus] K20D4
11	7505807CD1	g14603028	2.6E-110	[Homo sapiens] DC12 protein
12	7506413CD1	g15192151	1.1E-138	[Mus musculus] IIIG9 long form
14	1740413CD1	592769  FLJ10078	5.4E-96	[Homo sapiens][DNA-binding protein] Protein containing a PHD-finger, which are implicated in chromatin-mediated transcriptional regulation, and a SET domain, which are involved in chromatin organization
17	5402506CD1	317333  Y75B8A.12	1.4E-106	[Caenorhabditis elegans] Protein with weak similarity to C. elegans Y18H1A_67.M
24	7506697CD1	592769  FLJ10078	2.5E-105	[Homo sapiens][DNA-binding protein] Protein containing a PHD-finger, which are implicated in chromatin-mediated transcriptional regulation, and a SET domain, which are involved in chromatin organization

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
26	7506416CD1	g15192151	2.9E-162	[Mus musculus] IIG9 long form Danielson, P.E., Sautkulis, L.N., Foye, P.E., Hedlund, P.B. and Carson, M.J., "A novel mRNA expressed along brain ventricles" Gene Expr. Patterns (2001) In press
27	4823849CD1	g14039845	8.5E-52	[Homo sapiens] testes development-related NYD-SP18
29	7504597CD1	344862 SYT5	5.1E-11	[Homo sapiens] Synaptotagmin 5, a member of a family of calcium sensor proteins that regulate exocytosis of synaptic vesicles. Mirmics, K. et al. (2000) Neuron 28:53-67.
31	7506025CD1	g12002032	3.9E-92	[Homo sapiens] brain myo42 protein
32	7506102CD1	g14424787	1.3E-72	[Homo sapiens] natural killer cell transcript 4
		623570 NK4	1.8E-73	[Homo sapiens][Extracellular (excluding cell wall)] Natural killer cell transcript 4, protein with an RGD motif that may play a role in cell adhesion, expressed by lymphocytes and is upregulated in mitogen-activated T cells and IL2 treated natural killer cells. Dahl, C.A. et al. (1992) J. Immunol. 148:597-603.
35	2815375CD1	748886 TREX1	8.7E-295	[Homo sapiens] Protein with high similarity to three prime repair exonuclease 1 (mouse Trex1), which is a DNA-specific 3' exonuclease that acts preferentially on mispaired 3' termini and also degrades single and double-stranded DNA, member of the exonuclease family.
38	4913449CD1	g11528081	5.4E-48	[Mus musculus] stromal protein associated with thymii and lymph nodes short isoform. Flomerfelt, F. A. et al. (2000) Genes Immunol. 1:391-401.
39	7506136CD1	g1688307	1.8E-127	[Homo sapiens] cisplatin resistance associated alpha protein
		599286 FLJ20313	8.8E-18	[Homo sapiens] Protein with low similarity to the myotubularin family of dual specificity protein phosphatases.
40	7506225CD1	g14424787	1.3E-72	[Homo sapiens] natural killer cell transcript 4

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		623570 NK4	1.8E-73	[Homo sapiens][Extracellular (excluding cell wall)] Natural killer cell transcript 4, protein with an RGD motif that may play a role in cell adhesion, expressed by lymphocytes and is upregulated in mitogen-activated T cells and IL2 treated natural killer cells. Dahl, C. A. (1992) J. Immunol. 148:597-603.
41	7506227CD1	g14043143	2.4E-96	[Homo sapiens] muscle specific gene
45	045680CD1	247598 K06A9.1	1.5E-11	[Caenorhabditis elegans] Putative mucin, has strong similarity to H. sapiens MUC1 gene product [mucin 1, transmembrane].
49	55091643CD1	692116 FLJ20288	8.0e-24	[Homo sapiens] Protein containing fifteen ankyrin (Ank) repeats, which may mediate protein-protein interactions.
		434390 K1AA037	1.1e-23	[Homo sapiens] Protein containing twenty-three ankyrin (Ank) repeats, which may mediate protein-protein interactions.
		746899 Ank3	7.7E-23	[Rattus norvegicus][Anchor Protein][Basolateral plasma membrane; Cytoplasmic; Plasma membrane; Axon; Cytoskeletal] Ankyrin, an anchoring protein that links integral membrane proteins to the cytoskeleton, binds Na <sup>+</sup> /K <sup>+</sup> -ATPase. Zhou, D. et al. (1997) J. Cell. Biol. 136:621-631.
55	7509081CD1	742592 DKFZP434C245	3.9E-157	[Homo sapiens] Protein containing WD domains (WD-40 repeat), which may mediate protein-protein interactions.
		753721 WDR5	1.9E-38	[Homo sapiens] WD repeat domain 5, contains seven WD domains (WD-40 repeats), which likely mediate protein-protein interactions.



Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		339802 PAFAH1 B1	8.6E-27	[Homo sapiens][Structural protein; Hydrolase][Cytoplasmic] Platelet-activating factor acetylhydrolase (isoform 1b) alpha subunit (45kD), a noncatalytic subunit of a cytosolic heterotrimeric enzyme that inactivates platelet-activating factor; mutation of the gene causes lissencephaly and Miller-Dieker syndrome. Reiner, O. et al. (1993) Nature 364:717-721.
		587231 Pafah1b1	8.6E-27	[Mus musculus] [Hydrolase] [Cytoplasmic; Cytoskeletal; Centrosome/spindle pole body; Cilium] Platelet-activating factor acetylhydrolase (isoform 1b) beta 1 subunit, a heterotrimeric enzyme that inactivates platelet-activating factor, reduced expression disrupts neuronal migration and embryonic development. Hirotsune, S. et al. (1998) Nat. Genet. 19:333-339.
		711742 Pafah1b1	8.6E-27	[Rattus norvegicus][Hydrolase][Cytoplasmic; Centrosome/spindle pole body] Platelet-activating factor acetylhydrolase beta subunit (PAF-AH beta), part of a heterotrimeric enzyme that inactivates platelet-activating factor, involved in brain development. Watanabe, M. et al. (1998) Biochim. Biophys. Acta 1401:73-79.
56	7502450CD1	g12803841	8.9E-160	[Homo sapiens] Similar to retinoic acid induced 12.
57	7501405CD1	476011 LOC512719	7.6E-10	[Homo sapiens][Hydrolase;Protease (other than proteasomal)] Member of the trypsin family of serine proteases, contains an extracellular CUB domain, has moderate similarity to CIR (complement component C1r), mutations in the gene for which are associated with lupus erythematosus-like disease.
63	7503282CD1	336986 STK16	1.8E-11	[Homo sapiens][Protein kinase;Transferase] Serine threonine kinase 16, a myristoylated and palmitoylated protein kinase that may regulate transcription in response to signaling by transforming growth factor beta. Ligos, J. M. et al., Biochem. Biophys. Res. Commun. 249, 380-4 (1998).

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
		586841 Stk16	2.2E-11	[Mus musculus][Protein kinase;Transferase] Serine threonine kinase 16, a protein kinase that may be involved in cell growth and embryogenesis. Ligos, J. M. et al. (supra)
64	7503284CD1	336986 STK16	1.8E-11	[Homo sapiens][Protein kinase;Transferase] Serine threonine kinase 16, a myristoylated and palmitoylated protein kinase that may regulate transcription in response to signaling by transforming growth factor beta. Ligos, J. M. et al. (supra)
		586841 Stk16	2.2E-11	[Mus musculus][Protein kinase;Transferase] Serine threonine kinase 16, a protein kinase that may be involved in cell growth and embryogenesis. Ligos, J. M. et al. (supra)
68	7640560CD1	g817954	8.2E-173	[Mus musculus] DMR-N9. Mahadevan, M. S. et al., Hum. Mol. Genet. 2, 299-304 (1993)
		320574 Dm9	6.9E-174	[Mus musculus] Dystrophia myotonia-containing WD repeat, putative signaling protein; gene disruption may contribute to mental and testicular symptoms seen in myotonic dystrophy. Jansen, G. et al., Nat Genet 1, 261-6 (1992).
		347028 DMWD	6.5E-164	[Homo sapiens][Protein kinase;Transferase] Dystrophia myotonia-containing WD repeat, putative protein serine-threonine kinase; gene disruption may contribute to symptoms seen in myotonic dystrophy. Jansen, G. et al. (supra)
69	7506087CD1	g10280603	2.3E-172	[Homo sapiens] WDR10p-L. Gross, C. et al., DNA Cell Biol. 20, 41-52 (2001)
		753713 WDR10	1.8E-173	[Homo sapiens] Protein containing three WD domains (WD-40 repeats), which likely mediate protein-protein interactions. Gross, C. et al., DNA Cell Biol 20, 41-52. (2001).

Table 3

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	1393336CD1	126	S23 S77 S84		Eukaryotic cobalamin-binding proteins signature: K43-M95	PROFILESCAN
2	1431502CD1	109	S7 S15 S54		Natriuretic proteins BL00263: G47-Q64 PROTEIN ALU SUBFAMILY NUCLEAR SUBUNIT PHOSPHORYLATION J PD005149: R28-S78 e-value: 1.1e-10	BLIMPS_BLOCKS BLAST_PRODROM
3	2445220CD1	464	S4 S160 S270 S283 S308 S317 S370 S375 S419 T126 T253 T334 T346 T352 Y249	N163 N316 N341 N347 N391		
4	5504385CD1	273	S4 S25 S195 S211 S222 S244 S263 S268 T104			
5	6974948CD1	877	S85 S86 S101 S133 S134 S139 S144 S156 S171 S193 S198 S212 S220 S321 S373 S405 S408 S515 S552 S665 S693 S813 S867 T37 T183 T233 T251 T293	N94 N157 N402 N427 N525 N566 N635	GRAM domain: E239-K306 Cytosolic domain: M790-H877 Transmembrane domain: L767-W789 Non-cytosolic domain: M1-L766 YDR326C; MEMBRANE PROTEIN; DM04911 P43560 45-326: S122-L357 DM04911 P38800 411-680: S139-K350 DM04911 S59792 513-793: E123-K350 Leucine zipper pattern: L777-L798	HMMER_PFAM TMHMMER BLAST_DOMO MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
6	7501636CD1	756	S42 S87 S162 S166 S170 S361 S444 S559 S648 S670 S687 T18 T167 T311 T362 T457 T584 T597 T612 T627 T637	N96 N360		
7	2535717CD1	363	S45 S58 S81 S116 S138 S285 S293 S299 S308 S319 S348 S356 T34 T62 T109 T177 T230 T324	N12	AUTOANTIGEN ISLET CELL P69 ICA69 ANTIGEN DIABETES MELLITUS TYPE I PD011796: N12-C306	BLAST_PRODOM
8	6119548CD1	608	S29 S38 S42 S73 S137 S273 S289 S444 S514 S565 S576 S601 T64 T97 T122 T128 T327 T421 T448 T589	N142	Signal carboxyl-terminal PF00512: V472-L490 Cell attachment sequence: R35-D37	BLIMPS_PFAM MOTIFS
9	72263451CD1	424	S5 S30 S134 S189 S267 S311 S332 S348 S393 T99 T373 Y355 Y417		Leucine zipper pattern: L67-L88	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
10	7502640CD1	913	S9 S14 S31 S117 S171 S309 S338 S448 S634 S700 S709 S774 S790 S795 S823 S834 S862 T73 T85 T214 T275 T299 T320 T326 T424 T472 T542 T545 T572 T606 T627 T802 T816	N71 N360 N506 N877 N900		
11	7505807CD1	264	S112 S143 S194 S205 T236 T240	N156 N167		
12	7506413CD1	553	S29 S38 S42 S73 S137 S273 S289	N142	Signal carboxyl-terminal PF00512: V472-L490 Cell attachment sequence: R35-D37	BLIMPS_PFAM MOTIFS
13	1283631CD1	263	S26 S62 S123 S128 S181 S182 S198 S200 S251 S260 T134 T159 T228		COSMID K02F3 PD108053: G2-K157	BLAST_PRODOR

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
14	1740413CD1	1449	S12 S30 S79 S91 S136 S141 S145 S182 S198 S246 S484 S513 S546 S548 S598 S674 S786 S936 S1027 S1033 S1034 S1074 S1113 S1126 S1138 S1145 S1151 S1170 S1195 S1253 S1280 T55 T170 T192 T255 T446 T516 T555 T565 T594 T665 T682 T706 T758 T1177 T1213	N11 N134 N208 N261 N268 N426 N580 N821 N850 N1168 N1211 N1336	SET domain: Q277-C403	HMMER_PFAM
15	1951731CD1	400	S24 S100 S118 S140 S302 T20 T42 T276 T379 Y150	N28 N244 N300	Cell attachment sequence: R144-D146	MOTIFS
16	3741930CD1	226	S17 S48 S99 S145 S194 T88 T108 T162	N84 N184	Cell attachment sequence: R163-D165	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
17	5402506CD1	715	S17 S110 S117 S203 S301 S344 S369 S430 S516 S601 S710 T20 T32 T77 T94 T223 T402 T444 T451 T452 T473 T548 T552 T557 T585 T605 Y354 Y363 Y465	N106 N569 N592		MOTIFS
18	71081333CD1	364	S88 S184 S279 S290 T60 T147 T327 T334 T344 T354	N135	PHOSPHATE AMINOTRANSFERASE PD00040: D315-P322	BLIMPS_PRODUM
19	7503139CD1	152	S6 S39		Cytosolic domain: M1-K41 Transmembrane domain: G42-V64 Non-cytosolic domain: G65-D152 Signal cleavage: M12-G79	TMHMMER SPSCAN
20	7505836CD1	756	S16 S95 S175 S236 S495 S535 S580 S581 S598 S684 S706 S717 S732 T194 T211 T667	N14 N197 N499 N565 N576	ATP/GTP-binding site motif A (P-loop): A393-S400	MOTIFS
21	7505838CD1	120	S117 S118 T86	N98	ATP/GTP-binding site motif A (P-loop): G9-T16 PROTEIN ATPBINDING I72AA LONG MJ1559 MTH1068 AF0814 PD013279: V5-K120	MOTIFS BLAST_PRODUM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
22	7505872CD1	328	S5 S30 S51 S119 S149 S217 S303 S321 T108	N20	Leucine zipper pattern: L295-L316, L302-L323	MOTIFS
23	7506456CD1	214	S17 S48 S99 S133 S182 T88 T150	N84 N172	F25H2.12 PROTEIN PD142741: V24-L200 Cell attachment sequence: R151-D153	BLAST_PRODOM MOTIFS
24	7506697CD1	1442	S15 S23 S72 S84 S129 S134 S138 S175 S191 S239 S477 S506 S539 S541 S591 S667 S779 S929 S1020 S1026 S1027 S1067 S1106 S1119 S1131 S1138 S1144 S1163 S1188 S1246 S1273 T10 T48 T163 T185 T248 T439 T509 T548 T558 T587 T658 T675 T699 T751 T1170 T1206	N127 N201 N254 N261 N419 N573 N814 N843 N1161 N1204 N1329	SET domain: Q270-C396	HMMER_PFAM
25	7623472CD1	268	S2 S8 S68 S79 S99 S178 S188 S210 S215 T38 T72 T86 T123 T218 T256 Y108	N6 N136 N204 N247	Leucine zipper pattern: L224-L245	MOTIFS



Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
26	7506416CD1	588	S29 S38 S42 S73 S137 S273 S289 S424 S494 S545 S556 S581 T64 T97 T122 T128 T327 T428 T569	N142	Cell attachment sequence: R35-D37	MOTIFS
					Signal Cleavage: M46-Q94	SPSCAN
					CHROMOSOME XV READING FRAME ORF YOL071W PROTEIN C12B10.06C I PD032855: R57-G123	BLAST_PRODOM
27	4823849CD1	306	S14 S233 S255 S279 T2 T108 T142 T144 T178 T201 T246 Y131 Y148	N95 N116 N197 N244 N252	signal_cleavage: M1-A54	SPSCAN
28	4433922CD1	239	S89 S121 S197 T146 T208			
29	7504597CD1	49	S19 T15 T38			
30	7505987CD1	247	S19 S187 T99 T110 Y186		Cytosolic domain: M1-R140 Transmembrane domain: L141-A163 Non-cytosolic domain: L164-C247	TMHMMER
31	7506025CD1	418	S36 S41 S69 S73 S78 S125 S157 S160 S299 S312 S348 T89 T115 T258 T411		G-protein coupled receptors signature: S136-L152	MOTIFS

Table 3

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
32	7506102CD1	139	S54 S105 S137		NATURAL KILLER CELLS PROTEIN 4 PRECURSOR SIGNAL PD116715: E19-K139, MI-L31	BLAST_PRODOR
					Cell attachment sequence: R121-D123	MOTIFS
33	1333949CD1	295	S50 S153 S226 T67 T259 T269 Y243 Y267	N30 N39 N48 N93 N161 N211	Cytosolic domain: M1-L11 Transmembrane domain: L12-F31 Non-cytosolic domain: T32-S295	TMHMMER
					Cell attachment sequence: R114-D116	MOTIFS
34	7035533CD1	540	S36 S120 S153 S204 S303 S347 S464 S475 T49 T84 T125 T203 T402 T412 T497 T537	N462 N533		
35	2815375CD1	791	S9 S13 S31 S99 S154 S175 S182 S209 S224 S282 S287 S296 S300 S335 S458 S518 S625 S641 S731 T56 T462 T575 T640 T700 T716 T766 T781	N256 N510 N574	Cell attachment sequence: R284-D286	MOTIFS
36	2820152CD1	154	S15 S38 S47 S87 S91 S107 T65 Y125	N74 N127		

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
37	2959305CD1	957	S195 S241 S257 S300 S317 S345 S399 S418 S446 S460 S482 S484 S490 S498 S503 S538 S570 S720 S734 S789 S821 S826 S887 S944 T29 T42 T101 T431 T440 T524 T583 T595 T849 T932 Y647	N100 N296 N301 N438 N496 N580 N593 N914 N953	EG: EG0002.3 PROTEIN PD185691:M1-V240	BLAST_PRODUM
38	4913449CD1	340	S3 S14 S20 S123 S184 S193 S202 S203 S297 S309 S334 T47 T84 T149 T177 T272 T307 T329 T333			
39	7506136CD1	287	S23 S94 S188 T198 T204 T218 T277		CISPLATIN RESISTANCE ASSOCIATED PROTEIN ALPHA BETA PD038509: M21-L229	BLAST_PRODUM
					CISPLATIN RESISTANCE ASSOCIATED ALPHA PROTEIN PD117874: P230-P257	BLAST_PRODUM
40	7506225CD1	139	S54 S105 S137		NATURAL KILLER CELLS PROTEIN 4 PRECURSOR SIGNAL PD116715: E19-K139, M1-L31	BLAST_PRODUM
					Cell attachment sequence: R121-D123	MOTIFS

Table 3

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
41	7506227CD1	185	S104 S147 S158 S162 T99			
42	3144431CD1	164	S91 S153 S155 T28 T82 T105 T114	N51		
43	2633315CD1	577	S27 S244 S246 S274 S305 S395 S450 S455 S466 S471 S473 S475 S483 T73 T102 T132 T220 T289 T347 T513 Y32	N58 N258 N562		
44	3401751CD1	313	S56 S184 S211 S215 S217 T71 T104 T108 T229	N5 N100 N259	Cytosolic domain: Q313-Q313 Transmembrane domain: V290-I312 Non-cytosolic domain: M1-G289	TMHMMER
45	045680CD1	837	S220 S237 S244 S257 S263 S333 S369 S463 S513 S599 S605 S656 S668 S766 S819 S820 S832 T11 T79 T88 T100 T290 T301 T353 T676 T737 T738 T811	N218 N230 N800		
46	1503172CD1	195	S43 S47 S61 S74 S116 S174 T36 T84 T162	N139		

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
47	1818665CD1	1361	S192 S311 S320 S400 S478 S535 S571 S586 S588 S603 S656 S668 S771 S829 S849 S893 S899 S933 S943 S983 S990 S1045 S1057 S1068 S1139 S1232 S1244 S1271 S1297 S1326 S1358	N177 N240 N508 N551 N554 N754 N897	Cytosolic domain: L144-S195 Transmembrane domain: Y121-F143 Non-cytosolic domain: M1-L120	TMHMMER
			T64 T166 T378 T419 T474 T625 T663 T678 T791 T843 T995 T1010 T1081 T1087 T1089 T1266 T1313 T1346		Filaggrin signature: PR00487: A826-S868, S846-Q868, D976-N991	BLIMPS_PRINTS
					Cell attachment sequence: R380-D382	MOTIFS
48	3251352CD1	552	S48 S110 S344 T14 T26 T242 Y378		Leucine zipper pattern: L160-L181	MOTIFS
49	55091643CD1	345	S48 S119 S176 S224 S305 S328 T104	N166 N174	Ank repeat: N277-E309, F143-G175, G209-T241, S176-L208, S243-R276	HMMER_PFAM
					Ank repeat protein: PF00023: L181-L196, G210-D219	BLIMPS_PFAM
50	7500770CD1	30				

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
51	7506350CD1	685	S25 S118 S126 S127 S145 S156 S226 S255 S268 S332 S374 S447 S457 S459 S472 S478 S517 S522 S588 S608 S647 T151 T346 T400 T464 Y305	N585		
52	7508370CD1	104	S43 S47 S61 S74 T36 T84			
53	2894093CD1	672	S3 S22 S23 S113 S118 S142 S155 S239 S254 S258 S275 S329 S335 S343 S361 S426 S431 S504 S550 S556 S561 S566 S584 S630 S635 S641 T70 T105 T220 T225 T388 T398 T510 T528 Y459	N159 N218 N325 N333		

Table 3

SEQ ID NO:	Incye Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
54	7507335CD1	689	S33 S44 S99 S190 S226 S390 S406 S425 S557 S617 S671 S673 S677 S679 S681 T21 T52 T64 T225 T237 T474 T522 T567 T646 Y433 Y519	N665	PPR repeat: H257-V291, T574-S608	HMMER_PFAM
					PPR: pentatricopeptide repeat domain: H257-V291, Q333-L369, Y292-L332, T574-S608	HMMER_PFAM
55	7509081CD1	359	S162 S175 S182 S300 S330 T23 T84 T105 T120 T126 T132 T168 T191 T210 T218 T252 T337	N326	WD40 repeats: R134-D173, P50-V89, S176-D215, E260-K299, T218-D257, D8-H47, V92-A131	HMMER_PFAM
					WD domain, G-beta repeat: F137-D173, L11-H47, L221-D257, L263-K299, E95-A131, R53-V89, C179-D215	HMMER_PFAM
					Trp-Asp (WD) repeat proteins BL00678: S162-W172	BLIMPS_BLOCKS
					WD-40 repeat signature C150-D194, T65-V109, T23-F69, S108-F153	PROFILES SCAN
					Beta G-protein (transducin) PR00319: I160-K174, P197-W214	BLIMPS_PRINTS
					G-Protein Beta WD-40 repeat PR00320: I160-K174	BLIMPS_PRINTS
					Trp-Asp (WD) repeats signature: I202-V216	MOTIFS

Table 3

SEQ ID NO:	Incye Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
56	7502450CD1	316	S64 S93 S146 S259 N91 S277 T2 T11			
57	7501405CD1	60	S28			
58	7504528CD1	205	T166			
59	7509049CD1	203	S6 S98 S174 T154			
60	7509086CD1	161	S57			
61	7506914CD1	81	S37 T22	N19		
62	5606114CD1	214	S138 S163 S171 S191 S192 S209 T199	N127 N140 N180	Signal peptide: M34-A73	SPSCAN
63	7503282CD1	36			pfkB family of carbohydrate kinases signatures: A4-G34	PROFILESKAN
64	7503284CD1	64	S35 S43 T38		pfkB family of carbohydrate kinases signatures: A4-S60	PROFILESKAN
65	7510501CD1	142	S66 S91 S99 S119 S120 S137 T127	N55 N68 N108		
66	7500444CD1	42	T3 T14		Cytosolic domain: M1-Q19; Transmembrane domain: P20-L39; Non-cytosolic domain: M40-V42	TMHMMER
67	7510297CD1	36	S16			
68	7640560CD1	569	S208 S290 S318 S328 S330 S353 S357 S448 S465 S488 S492 T18 T102 T137 T203 T217 T257 T279 T317 T366 T501 Y96	N346 N356 N405 N481	WD domain, G-beta repeat: L254-S290, L213-N248, N142-N178, V296-S328, L526-G559, L352-D382, T99-D131	HMMER_PFAM



Table 3

SEQ ID NO:	Incye Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					WD40 repeats: S251-S290, T209-N248, D293-D382, L140-N178	HMMER_SMART
					Trp-Asp (WD) repeat proteins BL00678: T279-W289	BLIMPS_BLOCKS
					G-protein beta WD-40 repeat signature PR00320: F369-L383, I277-F291	BLIMPS_PRINTS
					BETA-TRANSDUCIN FAMILY Y TRP-ASP REPEATS DM00005[Q08274]230-294: T185-N248	BLAST_DOMO
					BETA-TRANSDUCIN FAMILY Y TRP-ASP REPEATS DM00005[Q08274]295-340: F249-D293	BLAST_DOMO
69	7506087CD1	433	S80 S105 S135 S158 S189 S257 T6 T179 T205 T410 Y57	N408	WD domain, G-beta repeat: R202-Q238, S73-N108, L5-D39, K116-Q156	HMMER_PFAM
					WD40 repeats: M1-D39, D42-N108, N110-Q156, S158-T197, D199-Q238	HMMER_SMART
					Copies of WD repeat: G43-S68, G200-Q238	HMMER_SMART
					Beta G-protein (transducin) signature PR00319: L49-G65, P220-Y237	BLIMPS_PRINTS

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
70/1393336CB1/ 1879	1-253, 1-258, 1-511, 1-624, 1-1879, 13-264, 210-825, 213-678, 334-561, 345-894, 610-1040, 655-950, 709-950, 770-1267, 851-1124, 851-1363, 857-1132, 989-1817, 1001-1262, 1001-1273, 1045-1332, 1108-1822, 1122-1603, 1155-1385, 1272-1807, 1327-1879
71/1431502CB1/ 1453	1-783, 1-804, 1-810, 1-821, 1-880, 1-886, 1-897, 104-1026, 105-1026, 203-1026, 302-1026, 385-1026, 392-595, 392-629, 717-1238, 718-989, 718-1016, 718-1130, 718-1222, 718-1230, 718-1236, 718-1282, 718-1365, 773-1223, 822-1315, 930-1361, 946-1436, 1094-1126, 1228-1453, 1404-1436
72/2445220CB1/ 2250	1-510, 79-334, 80-334, 98-365, 104-295, 113-774, 136-339, 138-470, 142-787, 143-388, 144-652, 154-408, 154-683, 176-446, 186-407, 186-712, 186-740, 195-988, 197-833, 200-418, 211-465, 212-428, 284-843, 422-630, 458-727, 473-735, 517-1136, 525-1069, 610-912, 731-1347, 820-1469, 847-1312, 850-1128, 961-1617, 1060-1623, 1086-1309, 1086-1524, 1095-1437, 1121-1454, 1127-1664, 1154-1377, 1160-1539, 1221-1475, 1230-1265, 1257-1545, 1279-1502, 1321-1634, 1334-1590, 1348-2209, 1369-1610, 1396-1664, 1396-1787, 1411-1684, 1438-2233, 1519-1786, 1519-1812, 1549-1835, 1597-1872, 1598-2237, 1609-2198, 1650-2237, 1725-2240, 1757-2005, 1761-2048, 1761-2052, 1761-2070, 1766-1952, 1775-2006, 1782-2250, 1789-2019, 1796-2249, 1796-2250, 1800-2213, 1803-2055, 1809-2248, 1917-2164, 1917-2243, 1917-2250, 2050-2247, 2077-2250, 2079-2250, 2084-2242
73/5504385CB1/ 1652	1-557, 41-334, 196-479, 232-479, 234-439, 244-502, 244-826, 247-738, 250-511, 278-675, 288-732, 453-826, 457-732, 567-812, 578-823, 668-1122, 676-997, 816-1124, 826-1220, 884-1093, 889-1122, 1178-1310, 1181-1652
74/6974948CB1/ 5821	1-514, 38-660, 149-414, 314-937, 628-1231, 641-867, 860-1463, 868-1153, 947-1343, 1180-1594, 1218-1876, 1287-1592, 1553-1824, 1620-1879, 1620-1993, 1696-2009, 1778-2341, 1796-2053, 1858-2114, 1927-2443, 1966-2667, 2111-2569, 2121-2414, 2181-2487, 2181-2630, 2209-2884, 2292-2732, 2306-2749, 2326-2971, 2375-2732, 2378-2718, 2395-3046, 2400-2742, 2433-2749, 2779-3174, 2779-3183, 2790-3183, 2794-3183, 2858-3183, 2924-3344, 3010-3589, 3068-3327, 3103-3360, 3147-3589, 3221-3580, 3221-3588, 3221-3590, 3261-3588, 3305-3556, 3342-3794, 3367-3589, 3378-3573, 3397-3515, 3397-3863, 3479-4126, 3616-4289, 3646-4171, 3682-3947, 3727-4374, 3884-4568, 4035-4498, 4092-4419, 4227-4568, 4396-4971, 4400-4649, 4400-4974, 4430-4868, 4844-5493, 4894-5268, 4894-5821, 4897-5388, 4921-5308, 4958-5233

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
75/7501636CBI/ 3063	1-472, 1-2878, 9-496, 9-639, 9-642, 15-233, 20-591, 22-185, 22-239, 22-313, 24-247, 24-280, 24-638, 28-624, 30-625, 32-652, 34-654, 39-702, 41-545, 51-204, 58-835, 67-267, 75-887, 92-374, 100-966, 131-999, 188-675, 268-868, 359-672, 395-997, 420-686, 430-642, 430-889, 473-768, 473-949, 473-972, 473-1098, 473-1101, 473-1111, 474-1156, 505-1162, 508-1124, 553-748, 561-745, 569-839, 573-1162, 578-1122, 602-888, 652-1118, 658-884, 658-885, 662-1282, 665-1053, 683-1210, 684-1210, 686-1282, 702-1326, 720-972, 720-1201, 730-1523, 751-1062, 780-1457, 786-1004, 812-1058, 850-1092, 850-1115, 858-1127, 924-1516, 935-1040, 936-1630, 941-1482, 941-1504, 968-1164, 974-1563, 987-1284, 989-1568, 1004-1061, 1049-1716, 1084-1190, 1110-1848, 1117-1683, 1118-1575, 1121-1448, 1121-1533, 1121-1591, 1121-1614, 1121-1625, 1121-1636, 1121-1650, 1121-1652, 1121-1653, 1121-1659, 1121-1663, 1121-1711, 1121-1721, 1121-1723, 1121-1749, 1121-1775, 1124-1326, 1124-1569, 1125-1765, 1128-1620, 1141-1640, 1149-1450, 1165-1798, 1174-1655, 1174-1841, 1175-1714, 1183-1927, 1198-1554, 1203-1460, 1205-2032, 1218-1917, 1222-1454, 1262-1444, 1283-1937, 1283-1955, 1294-1898, 1306-1594, 1320-1870, 1329-1779, 1356-1938, 1382-1757, 1397-1663, 1401-1897, 1401-1962, 1401-1983, 1401-1989, 1402-1885, 1402-1913, 1427-1668, 1433-2244, 1436-2057, 1457-2163, 1477-2063, 1488-2291, 1491-1794, 1526-1734, 1548-2000, 1553-2271, 1579-1983, 1584-1875, 1604-2266, 1613-1874, 1620-2265, 1662-1876, 1680-1772, 1704-2265, 1705-2320, 1722-1922, 1734-2106, 1793-2006, 1805-2307, 1821-2042, 1842-2092, 1845-2049, 1852-2217, 1860-2098, 1898-2175, 1914-2018, 1918-2510, 2012-2248, 2023-2279, 2023-2295, 2087-2320, 2123-2320, 2131-2408, 2138-2320, 2184-2233, 2199-2495, 2209-2451, 2229-2796, 2234-2320, 2247-2715, 2320-2542, 2320-2569, 2320-2578, 2320-2585, 2320-2588, 2320-2672, 2320-2799, 2320-2878, 2320-2951, 2323-2585, 2323-2601, 2324-2529, 2324-2589, 2324-2615, 2324-2840, 2329-2448, 2329-2517, 2337-2855, 2349-2890, 2359-2622, 2373-2657, 2392-2640, 2403-2642, 2411-2590, 2412-2701, 2414-2614, 2422-2596, 2453-2741, 2458-2711, 2466-2730, 2473-2770, 2484-2718, 2484-2765, 2486-2595, 2488-2745, 2492-2770, 2503-2765, 2514-3063, 2546-2813, 2546-2829, 2555-2803, 2555-2816, 2574-2868, 2629-2871, 2629-2876, 2632-2851, 2641-2965, 2741-2978, 2772-2864
76/2535717CBI/ 3100	1-3100, 200-564, 201-451, 201-589, 201-620, 201-750, 201-774, 201-795, 201-851, 201-3074, 202-674, 208-459, 266-1010, 310-1010, 435-1010, 517-1010, 561-734, 572-1246, 599-824, 669-1010, 676-960, 676-1010, 1011-1295, 1011-1424, 1011-1492, 1095-1706, 1108-1545, 1222-1724, 1243-1679, 1263-1585, 1294-1704, 1326-1659, 1350-1604, 1356-1558, 1375-1963, 1380-1990, 1400-1678, 1416-2257, 1439-1975, 1448-2053, 1458-2015, 1534-2054, 1544-1985, 1617-1896, 1617-2140, 1647-1933, 1758-1990, 1892-2249, 1905-2711, 2062-2326, 2351-2628, 2386-2631, 2386-2835, 2410-2577, 2410-2680, 2462-2974, 2490-2950, 2506-3048, 2574-3038, 2581-3057
77/6119548CBI/ 2210	1-253, 1-589, 18-612, 20-612, 27-612, 37-612, 38-544, 38-590, 538-852, 538-1161, 551-1111, 661-2201, 767-1392, 767-1402, 767-1472, 767-1473, 907-1164, 1033-1262, 1084-1322, 1176-1610, 1176-1765, 1283-1979, 1292-1368, 1297-1620, 1366-1621, 1379-1664, 1409-1614, 1409-1645, 1409-1913, 1410-1696, 1445-2126, 1484-2140, 1508-1963, 1752-2194, 1753-2076, 1770-2049, 1770-2060, 1785-2194, 1787-2210, 1813-2199, 1818-2091, 1818-2182, 1818-2198, 1821-2199, 1848-2200, 1977-2200, 1989-2155, 2058-2200

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
78/72263451CB1/ 2460	1-791, 25-2459, 110-708, 110-736, 110-781, 188-803, 188-907, 189-658, 189-815, 196-406, 329-819, 428-936, 459-996, 531-881, 536-815, 536-912, 554-813, 554-1042, 555-1052, 588-894, 671-1219, 701-1340, 741-1125, 763-1031, 769-1269, 783-1424, 891-1317, 903-1515, 905-1640, 912-1502, 973-1308, 986-1502, 1005-1521, 1005-1595, 1005-1622, 1005-1634, 1012-1587, 1013-1261, 1013-1497, 1013-1647, 1013-1785, 1070-1502, 1079-1502, 1094-1356, 1147-1420, 1186-1468, 1187-1484, 1236-1486, 1318-1874, 1351-2026, 1352-1648, 1354-1976, 1355-2001, 1371-1749, 1381-1647, 1381-1735, 1381-1824, 1381-1861, 1381-1893, 1381-1897, 1381-1926, 1381-1954, 1381-2013, 1381-2014, 1381-2020, 1381-2023, 1381-2135, 1381-2182, 1381-2207, 1381-2220, 1381-2274, 1384-1875, 1391-2016, 1430-1612, 1437-1539, 1450-2005, 1454-1883, 1455-2061, 1491-1964, 1493-1791, 1596-1867, 1608-2352, 1636-1962, 1643-1889, 1676-2250, 1680-1892, 1693-1885, 1693-2169, 1699-1935, 1718-1984, 1747-2459, 1769-2380, 1773-2364, 1794-2200, 1799-2079, 1805-2458, 1822-2435, 1849-2435, 1871-2181, 1871-2262, 1871-2460, 1884-2456, 1902-2456, 1907-2151, 1932-2436, 1933-2435, 1948-2459, 1950-2456, 1974-2444, 1995-2372, 1995-2405, 2002-2460, 2003-2460, 2005-2207, 2010-2359, 2020-2213, 2068-2371, 2073-2456, 2074-2273, 2088-2205, 2088-2335, 2090-2378, 2143-2414, 2152-2425, 2157-2442, 2167-2442, 2175-2442, 2176-2421, 2176-2428, 2176-2443, 2205-2443
79/7502640CB1/ 3496	1-516, 7-678, 29-332, 30-539, 34-295, 36-582, 47-293, 47-478, 57-742, 62-321, 62-519, 64-787, 65-705, 67-803, 74-857, 75-586, 77-374, 89-777, 200-954, 398-1008, 417-475, 673-1297, 753-1096, 754-1118, 754-1138, 754-1154, 754-1173, 754-1234, 754-1331, 758-1352, 763-2914, 828-1611, 849-1096, 864-1111, 864-1451, 929-1729, 941-1483, 941-1701, 954-1586, 1087-1536, 1178-1821, 1257-1506, 1294-1900, 1307-1606, 1317-1756, 1336-1908, 1347-1756, 1352-1869, 1361-1978, 1392-1900, 1403-1880, 1414-1668, 1421-1672, 1451-2036, 1451-2045, 1474-1890, 1477-1970, 1498-1863, 1516-2028, 1516-2085, 1550-1816, 1577-2129, 1578-2157, 1580-2094, 1584-2233, 1592-2131, 1599-1825, 1606-1738, 1611-2117, 1619-2105, 1639-2238, 1671-2007, 1686-1860, 1691-1984, 1703-2359, 1711-2293, 1728-1840, 1743-2247, 1743-2294, 1754-2123, 1770-2322, 1787-2329, 1797-2423, 1802-2258, 1811-2332, 1824-2056, 1825-2292, 1825-2356, 1846-2497, 1865-2392, 1867-2245, 1893-2450, 1915-2174, 1926-2330, 1926-2439, 1939-2450, 1966-2498, 1969-2487, 1973-2507, 1978-2416, 1978-2495, 2010-2280, 2015-2602, 2017-2591, 2029-2296, 2048-2314, 2057-2855, 2104-2368, 2104-2642, 2130-2603, 2203-2644, 2221-2684, 2242-2675, 2276-2470, 2423-2684, 2426-2684, 2486-2752, 2563-2852, 2594-3266, 2598-2859, 2646-2899, 2776-3154, 2782-3496, 2833-3496, 2838-3131, 2885-3163

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
80/7505807CBI/ 1334	1-435, 25-551, 26-372, 29-543, 36-351, 408-1253, 411-979, 430-942, 431-691, 431-958, 448-1135, 448-1145, 449-1040, 452-997, 456-864, 461-1052, 465-940, 477-868, 477-945, 477-1051, 477-1094, 477-1124, 477-1129, 477-1158, 477-1165, 480-975, 488-934, 493-1192, 494-950, 512-1077, 514-1040, 526-1095, 533-1221, 534-979, 549-1299, 565-725, 568-1282, 574-1149, 576-996, 582-1095, 586-1013, 586-1254, 599-1274, 601-1102, 602-949, 602-1062, 606-1254, 606-1279, 612-870, 612-887, 612-925, 612-982, 612-1106, 612-1151, 612-1152, 612-1176, 612-1198, 612-1230, 612-1238, 612-1250, 612-1255, 612-1262, 613-1201, 620-880, 635-1288, 635-1327, 636-1177, 638-1318, 640-1203, 643-1236, 643-1265, 646-1331, 647-968, 655-1086, 658-977, 659-957, 666-1296, 667-1281, 670-1327, 674-1269, 684-1313, 697-1316, 698-1282, 704-1296, 723-1296, 723-1299, 726-1209, 726-1248, 731-1288, 735-1021, 735-1334, 751-996, 756-1031, 758-1220, 758-1296, 758-1300, 761-1139, 765-1014, 769-1328, 770-1228, 771-958, 776-1254, 778-1328, 779-1130, 779-1299, 789-1300, 791-1115, 792-1328, 794-1270, 796-1329, 809-1292, 809-1328, 813-1299, 815-1073, 817-1084, 818-1133, 819-967, 820-1124, 824-1057, 824-1130, 824-1292, 828-1126, 829-1299, 830-1088, 837-1111, 837-1299, 837-1300, 846-1151, 851-1047, 852-1299, 855-1292, 860-1296, 865-1299, 868-1171, 869-1073, 877-1290, 877-1295, 878-1334
81/7506413CBI/ 2100	1-253, 1-589, 18-612, 20-612, 27-612, 37-612, 38-544, 38-590, 538-852, 538-1161, 551-1111, 661-2036, 767-1392, 767-1402, 767-1472, 767-1473, 907-1164, 1033-1262, 1084-1322, 1176-1610, 1176-1765, 1283-1837, 1292-1368, 1297-1620, 1366-1621, 1366-1986, 1409-1614, 1409-1645, 1409-1833, 1410-1696, 1445-1837, 1484-1837, 1508-1837, 1514-2050, 1626-2100, 1670-2029, 1752-1837, 1753-1837, 1770-1837, 1785-1837
82/1283631CBI/ 1301	1-483, 3-76, 217-443, 217-505, 217-591, 217-595, 217-616, 217-649, 217-671, 217-701, 217-710, 217-729, 217-733, 217-737, 217-754, 217-770, 217-776, 217-794, 217-795, 217-830, 217-861, 321-796, 351-701, 365-745, 367-1019, 368-694, 379-999, 429-731, 468-1076, 489-724, 489-1014, 505-1094, 510-1076, 525-1085, 646-1099, 663-1098, 702-1015, 709-1301, 725-1099, 727-1093, 727-1099, 791-1096, 861-1301, 864-1098, 917-978, 955-1219

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
83/1740413CB1/ 7783	1-523, 1-557, 1-7660, 308-618, 346-836, 493-1008, 516-999, 543-919, 557-945, 585-851, 638-1216, 802-1125, 869-1139, 919-1482, 943-1207, 968-1258, 968-1423, 968-1504, 968-1522, 968-1551, 968-1555, 1025-1561, 1087-1530, 1112-1775, 1163-1537, 1164-1613, 1182-1856, 1191-1613, 1213-1624, 1213-1874, 1243-1651, 1262-1717, 1266-1628, 1266-1869, 1278-1593, 1278-1941, 1290-1833, 1306-1609, 1310-1757, 1332-1838, 1343-1985, 1367-1947, 1377-2034, 1413-1674, 1469-2143, 1481-2042, 1498-2034, 1536-1869, 1543-2199, 1546-1810, 1546-1930, 1624-1930, 1624-2027, 1647-2140, 1658-1919, 1779-2046, 1779-2050, 1792-2437, 1803-2404, 1814-2232, 1857-2408, 1872-2561, 1873-2561, 1883-2299, 1921-2297, 1931-2624, 1984-2432, 2017-2720, 2092-2343, 2092-2577, 2098-2716, 2242-2949, 2272-2543, 2347-2705, 2347-2730, 2361-2654, 2368-2901, 2378-2731, 2404-2709, 2564-2955, 2588-3146, 2605-3213, 2625-2876, 2651-2928, 2664-3171, 2789-3261, 2895-3445, 2903-3486, 2934-3412, 2937-3355, 3027-3440, 3115-3627, 3185-3485, 3185-3698, 3187-3729, 3194-3805, 3194-3806, 3203-3461, 3203-3717, 3203-3719, 3203-3775, 3203-3783, 3221-3737, 3250-3719, 3275-3563, 3276-3555, 3299-3620, 3353-3619, 3358-3969, 3385-3719, 3393-3913, 3424-3839, 3441-4042, 3485-4064, 3486-3770, 3555-3808, 3561-4031, 3603-4027, 3630-4308, 3657-4175, 3660-4166, 3664-4222, 3680-4253, 3697-3973, 3711-4095, 3714-4307, 3715-4217, 3717-4449, 3724-4104, 3735-3829, 3759-4432, 3766-4449, 3769-4481, 3780-4082, 3836-4285, 3841-4360, 3873-4530, 3875-4371, 3903-4364, 3913-4242, 3934-4338, 3986-4286, 3992-4301, 3995-4643, 3997-4292, 4030-4567, 4036-4300, 4037-4340, 4070-4720, 4070-4741, 4073-4524, 4087-4386, 4089-4343, 4100-4691, 4109-4614, 4119-4771, 4142-4398, 4142-4638, 4142-4699, 4165-4479, 4170-4433, 4189-4459, 4190-4841, 4196-4775, 4197-4793, 4197-4833, 4200-4473, 4200-4620, 4206-4860, 4208-4506, 4211-4821, 4212-4485, 4212-4736, 4213-4825, 4225-4743, 4270-4851, 4278-4573, 4278-4574, 4278-4577, 4281-4879, 4295-4895, 4297-4568, 4302-4809, 4303-4575, 4312-4714, 4360-5141, 4393-5003, 4435-5012, 4437-5162, 4440-5030, 4456-4974, 4456-4982, 4462-4821, 4462-5089, 4469-5129, 4514-4811, 4524-4811, 4524-5070, 4525-4789, 4531-5083, 4536-5141, 4539-4834, 4555-5152, 4564-4820, 4566-5148, 4577-5150, 4579-5065, 4603-5176, 4613-4894, 4621-4880, 4625-4875, 4626-5033, 4634-5150, 4670-5123, 4670-5237, 4682-5113, 4683-5156, 4685-5157, 4687-5160, 4689-4987, 4692-5151, 4696-5157, 4698-5152, 4699-5158, 4700-5159, 4702-5113, 4703-5176, 4707-5160, 4709-5157, 4716-4917, 4716-4918,

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
	4716-4938, 4716-5194, 4718-5157, 4718-5171, 4719-5162, 4721-5155, 4723-5157, 4724-5181, 4726-4993, 4727-5176, 4729-5158, 4730-5065, 4734-4962, 4736-5156, 4738-5113, 4738-5157, 4739-5155, 4739-5157, 4741-5157, 4745-5154, 4745-5156, 4749-5113, 4753-5155, 4760-5157, 4762-5151, 4763-5122, 4765-4971, 4769-5154, 4775-5174, 4777-5146, 4778-5157, 4780-5158, 4782-5157, 4786-5146, 4789-5159, 4789-5160, 4789-5161, 4790-5157, 4791-5033, 4797-5161, 4798-5157, 4804-5006, 4804-5144, 4804-5181, 4810-5171, 4829-5160, 4830-5157, 4831-5036, 4831-5161, 4836-5160, 4839-5158, 4846-5158, 4861-5119, 4868-5160, 4869-5175, 4886-5113, 4890-5157, 4901-5169, 4913-5144, 4924-5144, 4924-5157, 4924-5248, 4940-5215, 4958-5144, 4961-5206, 4964-5194, 4974-5157, 4980-5167, 5001-5177, 5207-5438, 5219-5765, 5243-5705, 5280-5545, 5280-5550, 5280-5918, 5300-5596, 5334-5730, 5403-5649, 5415-5830, 5433-5697, 5447-5715, 5490-6047, 5503-5966, 5525-6218, 5528-6129, 5558-6155, 5614-6260, 5636-5872, 5636-5945, 5636-5960, 5645-6241, 5655-5898, 5655-6158, 5685-6197, 5694-6074, 5694-6075, 5700-6241, 5705-5893, 5743-6206, 5758-6216, 5766-6248, 5838-6276, 5891-6386, 5915-6200, 5922-6354, 5937-6206, 5946-6178, 6163-6417, 6163-6625, 6169-6398, 6202-6642, 6221-6387, 6274-6334, 6299-6493, 6335-6628, 6335-6696, 6392-6701, 6396-6677, 6397-6683, 6398-6682, 6398-6697, 6398-6699, 6398-6704, 6402-6452, 6412-6723, 6424-6724, 6431-6696, 6445-6669, 6445-6695, 6456-6838, 6473-6838, 6550-6834, 6567-6799, 6579-6833, 6704-6836, 6724-6832, 6724-6833, 6729-6846, 6737-6802, 6737-6840, 6737-6843, 6755-6833, 6779-6836, 6789-7080, 6789-7362, 6789-7479, 6956-7643, 6982-7647, 6994-7263, 7225-7352, 7225-7454, 7225-7781, 7225-7783, 7436-7714, 7558-7654, 7562-7747, 7562-7748
84/1951731CBI/ 1977	1-155, 1-547, 82-709, 477-1083, 477-1100, 477-1103, 484-1352, 486-1352, 489-1352, 661-1351, 662-1103, 692-1103, 910-1103, 920-1088, 920-1091, 920-1103, 920-1138, 920-1157, 920-1242, 920-1250, 920-1290, 920-1292, 920-1301, 920-1323, 920-1364, 920-1371, 920-1421, 920-1436, 920-1440, 920-1451, 920-1456, 920-1458, 920-1472, 920-1477, 920-1501, 920-1504, 920-1525, 920-1537, 920-1550, 920-1552, 920-1553, 920-1663, 920-1665, 920-1671, 920-1672, 920-1683, 920-1696, 920-1798, 940-1977, 949-1664, 949-1868, 954-1665, 982-1669, 987-1701, 988-1571, 1005-1727
85/3741930CBI/ 1376	1-300, 101-347, 104-197, 104-346, 105-403, 107-399, 107-516, 113-411, 115-414, 137-743, 142-654, 151-810, 163-451, 191-455, 236-534, 348-821, 374-670, 378-830, 391-670, 404-718, 457-880, 461-1157, 486-773, 530-1055, 532-775, 559-1145, 573-880, 614-1001, 655-1136, 693-1304, 767-1267, 768-880, 783-1376, 823-1350

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
86/5402506CB1/ 2635	1-318, 1-476, 1-603, 1-711, 16-246, 16-469, 16-504, 18-286, 23-629, 27-190, 30-321, 39-697, 43-308, 43-435, 46-546, 48-338, 48-543, 48-676, 49-145, 49-672, 49-674, 57-736, 58-800, 59-329, 74-266, 74-398, 74-698, 75-308, 302-863, 393-638, 529-869, 562-646, 605-1160, 607-804, 607-873, 619-909, 694-771, 759-1274, 770-1160, 800-1251, 891-1161, 902-1172, 951-1340, 980-1599, 985-1611, 1008-1236, 1009-1435, 1016-1511, 1016-1529, 1022-1637, 1060-1726, 1123-1579, 1167-1697, 1194-1813, 1213-1847, 1217-1429, 1217-1495, 1255-1570, 1281-1774, 1286-1776, 1298-1913, 1299-1832, 1299-1913, 1299-1919, 1310-1779, 1325-1943, 1342-1603, 1367-1631, 1380-1816, 1381-1945, 1392-1900, 1397-1473, 1428-1899, 1432-1887, 1432-1924, 1435-1864, 1436-1891, 1448-1960, 1459-1720, 1459-2034, 1461-2065, 1477-1725, 1481-2119, 1497-2055, 1500-2043, 1509-2080, 1510-1642, 1512-1762, 1534-1679, 1549-2101, 1638-2229, 1729-1979, 1758-2164, 1791-2060, 1896-2032, 2126-2418, 2139-2635, 2140-2609
87/1081333CB1/ 3414	1-271, 1-669, 28-244, 194-787, 238-441, 276-522, 276-691, 276-830, 340-502, 340-958, 653-993, 682-925, 682-958, 682-1110, 682-1180, 682-1195, 682-1199, 682-1224, 682-1230, 682-1235, 682-1254, 682-1288, 698-1383, 779-1235, 823-1465, 856-1305, 893-1480, 907-1409, 908-1529, 920-1182, 922-1372, 923-1631, 943-1135, 959-1212, 982-1250, 983-1572, 991-1452, 1036-1615, 1040-1688, 1082-1371, 1096-1562, 1099-1340, 1099-1437, 1099-1500, 1099-1509, 1099-1536, 1099-1539, 1099-1551, 1099-1560, 1099-1566, 1099-1590, 1099-1593, 1099-1599, 1099-1623, 1114-1416, 1120-1490, 1124-1742, 1143-1344, 1143-1376, 1143-1431, 1143-1618, 1143-1645, 1143-1670, 1143-1710, 1143-1768, 1143-1954, 1152-1701, 1154-1541, 1159-1709, 1160-1743, 1163-1686, 1240-1701, 1264-1517, 1271-1756, 1283-1539, 1285-1694, 1293-1839, 1302-1818, 1307-1537, 1312-1535, 1320-1921, 1339-1799, 1341-1938, 1352-1714, 1363-1929, 1364-1591, 1364-1603, 1364-1784, 1364-1786, 1364-1811, 1364-1828, 1364-1863, 1364-1881, 1364-1906, 1364-1908, 1364-1912, 1364-1927, 1364-1935, 1364-1948, 1364-2000, 1372-1869, 1380-1946, 1419-1787, 1420-1723, 1420-1726, 1436-1746, 1436-1929, 1438-1721, 1463-1891, 1468-2069, 1471-1929, 1503-2079, 1503-2154, 1514-1708, 1528-2101, 1534-2201, 1551-2221, 1555-1954, 1555-2042, 1557-1796,



Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
	1558-2094, 1581-2172, 1585-2176, 1586-2215, 1593-2019, 1597-2103, 1599-2204, 1605-2206, 1605-2236, 1609-1930, 1611-1930, 1620-2158, 1622-1843, 1622-2119, 1625-1914, 1626-2171, 1627-2060, 1635-2159, 1635-2168, 1645-2185, 1659-2022, 1660-2130, 1672-1862, 1683-2251, 1683-2262, 1691-2242, 1699-2247, 1699-2293, 1701-1929, 1707-2330, 1716-1866, 1717-2288, 1720-2075, 1724-1893, 1727-2164, 1731-1921, 1731-2237, 1735-2137, 1739-2243, 1755-2307, 1756-2279, 1758-2479, 1760-2027, 1760-2484, 1767-2014, 1776-2353, 1785-2032, 1785-2040, 1795-2204, 1797-2379, 1804-2277, 1805-2278, 1811-2059, 1812-2324, 1814-2310, 1834-2381, 1835-2381, 1850-2366, 1862-2415, 1868-2312, 1874-2108, 1876-2365, 1887-2293, 1887-2510, 1889-2294, 1889-2481, 1897-2139, 1903-2355, 1906-2521, 1915-2174, 1917-2358, 1918-2469, 1940-2511, 1942-2445, 1943-2188, 1945-2530, 1949-2520, 1950-2186, 1967-2460, 1983-2530, 1985-2252, 1985-2261, 1989-2477, 1998-2430, 1998-2521, 2004-2260, 2004-2514, 2006-2515, 2009-2517, 2014-2525, 2018-2521, 2019-2525, 2020-2511, 2050-2522, 2055-2525, 2059-2498, 2068-2520, 2069-2530, 2071-2299, 2071-2686, 2082-2476, 2083-2540, 2108-2371, 2139-2529, 2139-2530, 2188-2360, 2209-2525, 2232-2745, 2272-2549, 2284-2489, 2286-2483, 2286-2514, 2286-2525, 2288-2524, 2300-2430, 2314-2525, 2339-2527, 2346-2580, 2353-2609, 2391-2955, 2459-3046, 2472-2758, 2474-2693, 2474-2723, 2507-2987, 2524-2828, 2609-2664, 2622-2894, 2725-2967, 2735-3024, 2735-3031, 2748-2889, 2786-3371, 2823-3414, 2921-3098
88/7503139CB1/ 936	1-598, 17-503, 17-551, 17-648, 18-266, 18-540, 18-551, 18-648, 26-264, 48-637, 49-647, 63-708, 159-669, 181-596, 217-456, 346-651, 352-562, 352-579, 368-590, 423-936
89/7505836CB1/ 3123	1-347, 1-3123, 28-652, 29-380, 49-275, 97-639, 169-802, 200-603, 230-837, 325-881, 329-697, 410-945, 433-888, 434-903, 472-972, 479-1075, 514-694, 516-1068, 543-1111, 555-1172, 613-1220, 614-1172, 622-1260, 689-1238, 699-986, 787-1118, 796-1103, 813-1311, 819-1396, 829-1437, 848-1499, 858-1397, 858-1461, 881-1190, 889-1467, 930-1020, 954-1301, 985-1544, 1009-1459, 1015-1470, 1024-1799, 1033-1266, 1033-1484, 1076-1138, 1129-1839, 1132-1776, 1135-1351, 1139-1786, 1144-1843, 1145-1829, 1149-1298, 1172-1774, 1210-1497, 1210-1536, 1210-1634, 1210-1682, 1210-1691, 1210-1766, 1210-1837, 1210-1838, 1213-1838, 1218-1838, 1231-1838, 1257-1838, 1287-1955, 1296-1828, 1338-1753, 1355-1824, 1356-2287, 1358-1785, 1368-1838, 1378-1838, 1414-1828, 1425-1833, 1425-1895, 1428-1813, 1469-1829, 1487-2128, 1512-1829, 1522-1829, 1522-1838, 1526-1829, 1544-1748, 1570-1829, 1629-2184, 1660-1839, 1704-2200, 1704-2317, 1753-2345, 1766-1982, 1770-1991, 1770-2279, 1770-2629, 1868-2301, 1893-2092, 1989-2283, 2037-2290, 2054-2293, 2116-2789, 2121-2715, 2129-2407, 2161-2429, 2174-2951, 2206-2465, 2207-2671, 2211-2855, 2216-3055, 2288-2776, 2308-2921, 2321-2570, 2332-2947, 2332-2950, 2332-3048, 2332-3099, 2332-3110, 2335-3119, 2372-2648, 2390-3031, 2401-2864, 2405-2895, 2424-2710, 2444-2719, 2457-3114, 2486-2759, 2486-3061, 2503-3109, 2504-2626, 2506-2882, 2508-2839, 2521-3105, 2533-2801, 2542-3123, 2558-2952, 2605-2885, 2661-3111, 2678-3119, 2707-3119, 2723-3119, 2732-2927, 2734-3018, 2739-3119, 2740-2984, 2743-3007, 2815-3074, 2817-3117, 2845-3097, 2941-3119, 2963-3075

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
90/7505858CB1/ 616	1-140, 1-232, 1-234, 1-241, 1-253, 1-256, 1-280, 1-282, 1-288, 1-370, 1-381, 1-399, 1-616, 2-222, 2-284, 2-591, 5-279, 5-283, 5-332, 6-278, 7-265, 7-335, 8-243, 8-285, 9-286, 11-239, 11-377, 14-251, 14-286, 14-325, 16-268, 16-286, 17-268, 17-312, 18-237, 18-253, 18-292, 19-304, 20-309, 21-267, 22-257, 22-303, 22-316, 22-356, 24-303, 24-306, 25-303, 27-294, 33-285, 46-359, 50-285, 58-297, 80-333, 89-365, 91-353, 160-354, 181-442, 214-610, 250-367, 367-609, 406-603
91/7505872CB1/ 1834	1-237, 1-246, 1-258, 1-1821, 12-253, 27-616, 38-296, 149-777, 166-539, 195-401, 199-449, 210-435, 210-687, 237-493, 261-659, 312-736, 387-651, 387-659, 387-686, 407-877, 413-986, 413-1113, 436-875, 450-1071, 454-728, 454-1024, 456-739, 475-782, 475-1152, 475-1169, 475-1199, 476-752, 476-756, 477-682, 477-755, 477-1037, 477-1059, 477-1064, 477-1116, 477-1135, 477-1160, 477-1226, 482-1080, 483-754, 483-768, 483-770, 483-771, 483-986, 483-1161, 484-823, 487-1222, 533-644, 545-1267, 655-1117, 668-987, 709-937, 710-925, 712-1018, 718-1035, 738-1070, 751-963, 751-1238, 751-1272, 751-1278, 751-1280, 751-1295, 751-1305, 752-902, 760-1068, 761-1198, 765-1001, 768-1066, 771-1060, 773-932, 776-1024, 779-996, 785-1107, 792-1056, 792-1073, 792-1086, 801-970, 805-1192, 806-1135, 808-1200, 818-961, 839-1175, 840-1019, 846-1099, 853-1043, 858-1248, 859-1075, 873-1093, 879-1066, 893-1066, 897-1080, 900-1107, 933-1182, 938-1169, 960-1167, 960-1251, 963-1180, 1029-1184, 1052-1305, 1058-1496, 1059-1230, 1076-1185, 1168-1793, 1283-1494, 1310-1344, 1310-1574, 1310-1649, 1312-1480, 1312-1603, 1312-1814, 1313-1814, 1327-1559, 1332-1613, 1342-1571, 1353-1769, 1353-1770, 1361-1754, 1408-1814, 1421-1814, 1437-1814, 1455-1814, 1455-1832, 1468-1496, 1476-1496, 1493-1833, 1496-1814, 1497-1834, 1635-1834, 1637-1828, 1731-1814, 1732-1814, 1733-1814
92/7506456CB1/ 1340	1-300, 101-347, 104-197, 104-346, 105-403, 107-399, 107-516, 113-411, 115-414, 137-707, 142-618, 151-774, 163-451, 191-455, 236-523, 348-785, 374-634, 378-794, 391-634, 400-844, 404-682, 457-844, 461-1121, 486-737, 521-680, 522-1179, 523-1109, 524-739, 524-1019, 537-844, 555-845, 578-965, 619-1100, 657-1268, 731-1316, 732-844, 747-1340, 787-1314

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
93/7506697CBI/ 7706	1-560, 5-998, 5-5100, 207-560, 238-268, 373-891, 418-717, 466-842, 493-855, 508-774, 561-964, 561-1139, 692-802, 725-1048, 792-1062, 842-1405, 866-1130, 891-1181, 891-1346, 891-1364, 891-1427, 891-1445, 891-1474, 891-1478, 948-1484, 1010-1453, 1035-1698, 1086-1460, 1087-1536, 1105-1779, 1114-1536, 1136-1547, 1136-1797, 1166-1574, 1185-1640, 1189-1551, 1189-1792, 1201-1864, 1206-1456, 1213-1756, 1229-1532, 1233-1680, 1255-1761, 1266-1908, 1290-1870, 1300-1957, 1336-1597, 1382-1887, 1392-2066, 1404-1965, 1421-1957, 1466-2122, 1469-1733, 1469-1853, 1547-1853, 1547-1950, 1571-2063, 1581-1842, 1666-1831, 1694-1824, 1715-2360, 1726-2327, 1737-2155, 1780-2331, 1795-2484, 1796-2484, 1806-2222, 1822-1888, 1822-2047, 1844-2220, 1854-2547, 1907-2355, 1940-2643, 2015-2266, 2015-2500, 2021-2639, 2145-2364, 2165-2872, 2195-2466, 2270-2628, 2270-2653, 2284-2577, 2291-2824, 2301-2654, 2327-2632, 2487-2878, 2511-3069, 2528-3136, 2548-2799, 2574-2851, 2587-3094, 2712-3184, 2818-3093, 2818-3368, 2826-3409, 2857-3335, 2860-3278, 2950-3363, 3038-3550, 3108-3621, 3110-3652, 3117-3728, 3117-3729, 3126-3384, 3126-3408, 3126-3526, 3126-3640, 3126-3642, 3126-3698, 3126-3706, 3144-3660, 3173-3642, 3198-3486, 3199-3478, 3262-3418, 3281-3892, 3308-3642, 3316-3836, 3347-3762, 3364-3965, 3373-3498, 3408-3987, 3478-3731, 3484-3954, 3526-3950, 3553-4231, 3580-4098, 3583-4089, 3587-4145, 3596-3878, 3603-4176, 3620-3896, 3634-4018, 3637-4230, 3638-4140, 3640-4372, 3644-3816, 3655-3881, 3658-3752, 3682-4355, 3689-4372, 3692-4404, 3694-3951, 3703-4005, 3710-3955, 3731-4028, 3742-3997, 3759-4208, 3764-3900, 3764-4283, 3796-4453, 3798-4294, 3826-4287, 3836-4165, 3848-4007, 3857-4261, 3909-4209, 3915-4017, 3915-4224, 3918-4566, 3920-4215, 3959-4223, 3960-4263, 3993-4664, 3996-4347, 4010-4309, 4023-4614, 4032-4537, 4042-4694, 4065-4561, 4065-4622, 4088-4402, 4112-4382, 4113-4764, 4119-4698, 4120-4716, 4120-4756, 4123-4543, 4129-4783, 4134-4744, 4135-4659, 4136-4748, 4148-4666, 4193-4774, 4201-4496, 4201-4497, 4201-4500, 4204-4802, 4218-4818, 4220-4491, 4225-4732, 4226-4498, 4235-4637, 4240-4702, 4283-5064, 4296-4640, 4316-4926, 4358-4935, 4358-4950, 4360-4539, 4360-5085, 4363-4637, 4363-4953, 4379-4897, 4379-4905, 4385-4744, 4385-5012, 4386-4674, 4392-5052, 4437-4734, 4447-4734, 4447-4993, 4448-4712, 4454-5006, 4459-5064, 4462-4757, 4478-5075, 4487-4743, 4489-5071, 4493-5051, 4500-5073, 4502-4783, 4502-4988, 4526-5099, 4536-4817, 4544-4803, 4548-4798, 4549-4956, 4557-5073, 4571-5021, 4587-4927, 4593-5046, 4593-5101,

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
	4605-5036, 4606-5079, 4608-5080, 4610-5083, 4612-4910, 4615-5074, 4619-5080, 4621-5075, 4622-5081, 4623-5082, 4625-5036, 4626-5099, 4630-5083, 4632-5080, 4639-4840, 4639-4841, 4639-4861, 4639-5117, 4641-5080, 4641-5094, 4642-5085, 4644-5078, 4646-5080, 4647-5104, 4649-4916, 4650-5099, 4652-5081, 4653-4988, 4657-4885, 4659-5079, 4661-5036, 4661-5080, 4662-5078, 4662-5080, 4664-5080, 4668-5077, 4668-5079, 4672-5036, 4676-5078, 4683-5080, 4685-5074, 4686-5045, 4688-4894, 4692-5077, 4698-5097, 4700-5069, 4701-5080, 4703-5081, 4705-5080, 4709-5069, 4712-5082, 4712-5083, 4712-5084, 4713-5080, 4714-4956, 4720-5084, 4721-5080, 4727-4929, 4727-5067, 4727-5104, 4733-5094, 4752-5083, 4753-5080, 4754-4959, 4754-5084, 4759-5083, 4762-5081, 4769-5081, 4784-5042, 4791-5083, 4792-5098, 4809-5036, 4813-5080, 4824-5092, 4836-5067, 4847-5067, 4847-5080, 4847-5171, 4863-5138, 4881-5067, 4884-5129, 4887-5117, 4897-5080, 4903-5090, 4924-5100, 4968-5098, 5130-5361, 5142-5688, 5166-5628, 5203-5468, 5203-5473, 5203-5841, 5223-5519, 5257-5653, 5326-5572, 5338-5753, 5356-5620, 5370-5638, 5413-5970, 5426-5889, 5448-6141, 5451-6052, 5481-6078, 5537-6183, 5559-5795, 5559-5868, 5559-5883, 5568-6164, 5578-5821, 5578-6081, 5608-6120, 5617-5997, 5617-5998, 5623-6164, 5628-5816, 5666-6129, 5681-6139, 5689-6171, 5761-6199, 5814-6309, 5838-6123, 5845-6277, 5860-6129, 5869-6101, 6086-6340, 6086-6348, 6092-6321, 6125-6565, 6144-6310, 6197-6257, 6222-6416, 6258-6551, 6258-6619, 6315-6624, 6319-6600, 6320-6606, 6321-6605, 6321-6620, 6321-6622, 6321-6627, 6325-6375, 6335-6646, 6337-6619, 6347-6647, 6354-6619, 6368-6592, 6368-6618, 6379-6761, 6396-6761, 6473-6757, 6490-6722, 6502-6756, 6627-6759, 6647-6755, 6647-6756, 6652-6769, 6660-6725, 6660-6763, 6660-6766, 6678-6756, 6702-6759, 6712-7003, 6712-7285, 6712-7402, 6879-7566, 6905-7570, 6917-7186, 7148-7275, 7148-7377, 7148-7704, 7148-7706, 7359-7637, 7481-7577, 7485-7670, 7485-7671
94/7623472CBI/ 1663	1-280, 1-493, 6-190, 10-210, 10-279, 12-300, 14-448, 33-56, 76-347, 76-753, 142-439, 151-815, 189-758, 253-509, 253-781, 338-590, 341-885, 377-639, 420-742, 447-1076, 482-1018, 582-879, 602-1146, 635-1175, 671-927, 713-1299, 719-954, 750-965, 767-1002, 926-1213, 945-1431, 955-1611, 959-1515, 987-1220, 992-1614, 1004-1278, 1073-1335, 1102-1649, 1200-1503, 1230-1452, 1230-1616, 1262-1564, 1262-1643, 1262-1662, 1302-1654, 1309-1608, 1387-1660, 1388-1647, 1388-1648, 1388-1650, 1388-1663, 1402-1662
95/7506416CBI/ 2144	1-253, 1-589, 18-612, 20-612, 27-612, 37-612, 38-544, 38-590, 45-79, 538-852, 538-1161, 551-1111, 661-2144, 767-1253, 767-1392, 767-1402, 767-1472, 767-1473, 907-1164, 1033-1262, 1084-1322, 1175-1721, 1176-1288, 1176-1721, 1176-1729, 1292-1368, 1322-1624, 1482-1550, 1482-1554, 1482-1560, 1482-1561, 1482-1585, 1482-1604, 1482-1636, 1482-1705, 1482-1773, 1482-1777, 1482-1903, 1482-1919, 1482-2080, 1485-2066, 1494-1571, 1566-1777, 1569-1777, 1610-1777, 1644-1777, 1692-2134, 1693-2016, 1710-1989, 1710-2000, 1725-2134, 1727-2141, 1753-2139, 1758-2031, 1758-2122, 1758-2138, 1761-2139, 1778-1994, 1778-2001, 1778-2072, 1778-2074, 1788-2140, 1917-2140, 1929-2095, 1998-2140

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
96/4823849CB1/ 2117	1-363, 1-405, 3-642, 83-124, 83-196, 89-196, 230-501, 230-757, 271-876, 371-1060, 403-1076, 515-1198, 533-1082, 534-1141, 538-1174, 564-1195, 566-1245, 587-1292, 589-1200, 591-811, 591-1234, 598-1279, 598-1310, 601-1008, 611-1287, 612-1199, 630-1236, 643-1289, 652-1262, 657-1318, 659-1254, 669-1261, 680-1261, 695-1334, 709-1287, 714-1402, 729-1323, 747-1398, 765-1345, 771-1473, 818-1350, 831-1472, 833-1477, 834-1496, 898-1550, 913-1158, 934-1524, 1003-1579, 1058-1790, 1077-1798, 1112-1334, 1124-1755, 1146-1865, 1194-1951, 1207-1862, 1226-1797, 1263-1998, 1281-1961, 1351-1662, 1373-2074, 1542-2117, 1824-1937
97/4433922CB1/ 1713	1-440, 1-529, 1-570, 1-611, 20-642, 132-409, 158-725, 306-947, 319-961, 329-474, 329-867, 352-995, 396-663, 400-811, 467-823, 495-1057, 525-1065, 540-974, 541-1015, 545-1240, 549-691, 553-1204, 576-1195, 821-1375, 822-1110, 842-1058, 876-1138, 888-1135, 888-1144, 889-1312, 899-1196, 921-1383, 969-1197, 1247-1478, 1289-1538, 1478-1713
98/7504597CB1/ 1297	1-268, 1-1297, 16-286, 30-278, 151-839, 151-953, 151-996, 151-1000, 156-286, 283-809, 285-632, 285-1063, 288-443, 307-762, 314-1252, 321-1077, 338-985, 341-1252, 343-1252, 348-1077, 351-1072, 354-530, 368-976, 377-1252, 384-753, 390-888, 395-1002, 396-648, 401-949, 410-799, 416-1252, 426-1043, 427-1072, 434-1072, 435-1252, 444-917, 445-591, 458-1083, 469-1252, 471-1252, 480-1252, 481-1074, 486-1014, 499-762, 499-830, 514-1072, 517-1072, 521-1077, 543-1076, 555-1249, 562-1072, 567-681, 569-1072, 578-833, 579-1038, 590-1072, 594-1072, 598-1072, 599-869, 603-1072, 610-803, 610-1035, 612-1070, 612-1072, 615-865, 620-1072, 620-1076, 634-796, 635-1062, 637-1062, 637-1076, 638-1072, 640-1072, 646-1076, 647-1077, 649-927, 667-821, 670-1072, 683-1072, 687-1072, 689-836, 704-1062, 712-1075, 721-872, 725-971, 741-1039, 747-1076, 748-1056, 756-1038, 761-1038, 762-920, 820-1063, 861-1045, 890-1072, 1073-1297
99/7505987CB1/ 885	1-224, 1-250, 2-885, 4-267, 14-265, 14-607, 16-160, 23-304, 27-353, 44-287, 45-300, 62-267, 76-446, 78-348, 125-267, 128-464, 161-422, 267-490, 267-871, 274-795, 274-812, 275-452, 278-452, 297-572, 298-542, 305-431, 305-554, 305-559, 307-827, 319-821, 335-864, 343-630, 345-769, 356-642, 358-879, 363-579, 376-879, 387-882, 400-880, 409-646, 409-843, 409-873, 420-885, 425-885, 435-879, 449-884, 456-879, 458-749, 467-885, 495-880, 504-885, 523-881, 527-731, 529-821, 534-796, 536-665, 561-656, 570-885, 573-636, 597-821, 648-875, 648-881, 648-885, 649-885
100/7506025CB1/ 1652	1-301, 1-374, 1-377, 1-389, 1-410, 1-417, 1-439, 1-459, 1-484, 1-487, 1-490, 1-527, 1-561, 1-566, 1-581, 1-684, 1-1638, 16-198, 16-260, 16-261, 16-576, 16-618, 17-440, 18-261, 18-262, 19-521, 34-634, 41-594, 58-303, 74-484, 102-772, 114-665, 145-446, 145-471, 158-711, 189-819, 222-465, 279-791, 355-608, 355-627, 384-681, 408-649, 409-872, 421-1014, 456-1083, 479-1366, 486-923, 593-875, 601-803, 691-939, 702-1014, 714-990, 758-1014, 763-1014, 890-1165, 1011-1533, 1011-1638, 1012-1565, 1012-1603, 1014-1554, 1037-1185, 1048-1612, 1057-1501, 1060-1520, 1060-1570, 1079-1339, 1079-1603, 1091-1279, 1096-1507, 1125-1608, 1132-1607, 1142-1608, 1143-1426, 1171-1608, 1181-1609, 1198-1570, 1203-1606, 1207-1398, 1207-1402, 1207-1407, 1207-1433, 1207-1434, 1221-1473, 1222-1485, 1230-1499, 1246-1465, 1260-1490, 1291-1491, 1308-1602, 1312-1620, 1331-1480, 1334-1577, 1354-1613, 1421-1652, 1537-1608

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
101/7506102CB1/ 1022	1-180, 1-1022, 37-181, 44-181, 55-181, 64-181, 131-213, 239-303, 368-454, 371-595, 371-684, 371-719, 371-854, 371-1002, 376-458, 413-727, 458-900, 459-636, 459-658, 459-660, 459-662, 459-669, 459-672, 459-681, 459-685, 459-686, 459-696, 459-709, 459-712, 459-725, 459-726, 459-728, 459-737, 459-739, 459-747, 459-750, 459-760, 459-784, 459-938, 459-945, 459-957, 459-959, 459-1018, 459-1021, 460-691, 460-749, 461-1005, 462-728, 462-847, 465-1022, 466-999, 468-750, 468-764, 474-741, 475-573, 475-578, 477-758, 478-702, 479-742, 481-608, 481-756, 481-803, 487-686, 489-671, 489-699, 489-739, 491-714, 491-1003, 492-598, 492-748, 492-754, 493-723, 493-799, 494-754, 498-703, 498-755, 499-954, 502-744, 502-784, 502-907, 503-716, 504-744, 504-834, 505-799, 505-1020, 506-999, 511-1003, 512-775, 512-785, 512-999, 514-889, 515-786, 516-759, 517-748, 519-764, 523-889, 526-747, 526-848, 526-922, 526-979, 532-663, 536-726, 536-747, 536-822, 536-948, 536-1005, 541-882, 542-738, 542-752, 544-997, 544-999, 548-1005, 550-1022, 551-816, 553-816, 553-999, 553-1020, 553-1021, 553-1022, 556-988, 556-1016, 556-1019, 557-1022, 558-1021, 559-818,
	559-1005, 561-770, 563-864, 564-846, 565-898, 571-901, 571-916, 572-998, 573-821, 575-823, 579-703, 584-837, 584-985, 584-1022, 585-1002, 588-791, 591-997, 593-833, 595-803, 595-999, 596-1019, 599-1022, 600-1014, 604-839, 605-820, 608-812, 608-849, 608-999, 610-816, 612-1022, 617-996, 623-880, 623-1002, 625-998, 625-1022, 626-1011, 627-878, 627-999, 628-829, 628-873, 631-843, 631-844, 631-860, 631-999, 633-761, 633-1016, 634-1017, 638-835, 638-1018, 642-878, 643-1002, 644-1021, 646-895, 649-966, 650-917, 651-910, 654-908, 658-1004, 662-879, 664-860, 664-861, 668-1002, 669-935, 669-1016, 677-974, 685-950, 691-891, 692-1002, 696-930, 702-956, 702-964, 722-868, 726-999, 727-1005, 737-998, 737-1005, 741-1016, 742-963, 752-997, 755-1022, 758-1019, 760-1003, 764-1022, 770-980, 777-995, 777-999, 790-999, 798-959, 810-1019, 818-1019, 824-1009, 838-991, 840-1022, 857-973, 858-1020, 859-1015, 883-1019, 895-957, 895-1000, 900-993, 900-1006, 907-1002
102/1333949CB1/ 2046	1-214, 1-664, 276-986, 288-986, 622-1582, 634-1589, 1289-2046
103/7035533CB1/ 2355	1-579, 104-2346, 546-706, 546-826, 638-901, 708-1319, 709-979, 709-1197, 709-1213, 709-1270, 709-1319, 709-1331, 713-1320, 760-1304, 775-1047, 777-1328, 842-1540, 847-1026, 933-1145, 933-1218, 979-1541, 999-1630, 1031-1052, 1051-1285, 1051-1678, 1053-1579, 1083-1689, 1108-1719, 1162-1699, 1187-1848, 1244-1548, 1248-1810, 1256-1709, 1321-1952, 1372-1535, 1373-1628, 1373-1657, 1375-1993, 1442-1629, 1451-2017, 1455-1852, 1457-1989, 1473-1745, 1496-1750, 1497-1782, 1523-2083, 1540-1811, 1543-1878, 1545-1829, 1576-2118, 1581-2318, 1590-2238, 1621-2282, 1623-2274, 1634-1912, 1635-2274, 1636-2267, 1646-2271, 1648-2271, 1660-2321, 1660-2346, 1667-2291, 1689-2102, 1689-2110, 1697-1901, 1697-2207, 1709-2338, 1718-2346, 1768-2215, 1862-2346, 1876-2103, 1880-2346, 1883-2335, 1888-2355, 1889-2346, 1890-2355, 1896-2344, 1920-2251, 1931-2346, 1982-2346, 2094-2340, 2114-2207

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
103/7035533CB1/ 2355	1-702, 69-850, 85-504, 490-792, 490-924, 490-966, 490-1031, 491-603, 491-777, 492-1130, 607-1266, 696-1092, 699-1309, 752-1263, 801-1263, 842-1012, 847-1098, 849-1390, 890-1507, 959-1345, 972-1217, 972-1507, 982-1268, 997-1377, 998-1549, 1008-1201, 1025-1488, 1066-1487, 1136-1683, 1140-1695, 1146-1657, 1260-1847, 1270-1837, 1286-1897, 1300-1894, 1333-1838, 1366-1858, 1382-1905, 1415-2125, 1457-1914, 1470-1847, 1533-2108, 1535-2029, 1543-1810, 1544-2153, 1591-1726, 1625-1889, 1643-2022, 1689-2208, 1698-2309, 1768-2158, 1779-2040, 1789-2442, 1789-2444, 1795-2036, 1795-2266, 1801-2498, 1810-2435, 1824-2365, 1866-2270, 1869-2120, 1929-2453, 1949-2564, 2001-2533, 2027-2573, 2038-2252, 2038-2315, 2039-2480, 2039-2487, 2039-2537, 2039-2539, 2039-2559, 2041-2343, 2063-2300, 2063-2367, 2142-2537, 2154-2566, 2166-2536, 2167-2559, 2209-2570, 2210-2491, 2218-2528, 2223-2548, 2405-2573
105/2820152CB1/ 1488	1-165, 1-243, 1-416, 1-427, 1-433, 1-434, 1-520, 2-519, 3-245, 14-246, 45-715, 49-610, 72-415, 72-671, 86-581, 88-327, 89-676, 100-302, 100-310, 113-405, 158-475, 205-683, 289-802, 342-591, 347-547, 347-559, 347-918, 358-581, 411-605, 438-973, 469-773, 513-813, 536-822, 556-833, 556-1066, 560-845, 572-1077, 573-1089, 600-907, 709-948, 714-1278, 725-1007, 729-902, 740-998, 740-1011, 760-961, 760-980, 803-1282, 814-1051, 815-1282, 845-1127, 919-1107, 934-1488, 997-1483, 998-1460, 1033-1340, 1033-1476, 1051-1456, 1054-1465, 1095-1344, 1201-1463, 1279-1465
106/2959305CB1/ 4486	1-669, 1-679, 254-765, 390-835, 390-889, 390-923, 390-932, 390-935, 390-947, 501-1193, 546-888, 647-1196, 755-947, 881-1444, 1236-1724, 1292-1565, 1292-1815, 1301-1923, 1416-1850, 1479-1751, 1672-2293, 1672-2342, 1735-2203, 1764-2016, 1823-2415, 1858-2063, 1867-2499, 1867-2532, 1946-2252, 1950-2111, 2038-2661, 2177-2782, 2204-2468, 2269-2516, 2269-2872, 2277-2952, 2303-2892, 2349-2949, 2519-2753, 2594-2975, 2606-2909, 2612-2862, 2612-3053, 2617-3142, 2653-2894, 2666-3236, 3001-3462, 3035-3503, 3065-3634, 3078-3397, 3153-3402, 3153-3614, 3197-3617, 3198-3526, 3208-3773, 3300-3558, 3306-3575, 3338-3827, 3397-3805, 3407-3643, 3408-3636, 3408-3655, 3408-3910, 3408-3929, 3414-3550, 3479-3759, 3486-3722, 3568-3861, 3643-3942, 3643-4072, 3705-4174, 3736-4288, 3774-4372, 3813-4403, 3839-4128, 3840-4103, 3840-4354, 3842-4138, 3858-4143, 3868-4125, 3868-4147, 3878-4181, 3887-4185, 3890-4477, 3908-4143, 3916-4208, 3921-4171, 3921-4182, 3924-4226, 3935-4194, 3943-4456, 3947-4482, 3983-4416, 3989-4278, 3990-4424, 3996-4263, 3997-4326, 4003-4276, 4015-4277, 4175-4409, 4178-4409, 4220-4454, 4309-4486
107/4913449CB1/ 1223	1-247, 69-100, 69-255, 69-269, 69-352, 69-367, 71-255, 71-270, 71-287, 71-370, 71-371, 73-317, 73-334, 73-343, 73-1223, 79-306, 81-306, 88-373, 91-335, 94-456, 98-456, 163-737, 163-738, 497-1080, 505-1145, 628-900, 637-885, 681-884, 681-922, 681-1139, 681-1182, 681-1190, 681-1201, 854-1221
108/7506136CB1/ 1188	1-277, 1-444, 3-230, 5-1188, 22-527, 22-603, 30-492, 110-245, 110-383, 110-598, 110-608, 110-626, 110-702, 110-806, 110-893, 112-652, 156-1050, 157-669, 163-1055, 305-437, 324-864, 389-1040, 439-1009, 541-716, 544-975, 549-1037, 550-811, 601-875

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
109/7506225CB1/ 745	1-234, 1-323, 1-358, 1-493, 3-745, 10-482, 13-76, 13-97, 97-538, 98-275, 98-308, 98-320, 98-324, 98-325, 98-335, 98-364, 98-365, 98-367, 98-376, 98-378, 98-386, 98-389, 98-396, 98-399, 98-423, 99-330, 99-388, 100-491, 101-367, 101-486, 107-389, 107-403, 109-477, 113-380, 114-212, 116-397, 117-341, 118-381, 120-247, 120-395, 120-442, 126-325, 128-310, 128-338, 128-378, 131-237, 131-393, 132-362, 132-438, 133-393, 137-342, 137-394, 141-383, 141-423, 143-383, 143-473, 144-438, 150-395, 151-414, 154-425, 155-398, 158-403, 165-386, 165-487, 175-365, 175-386, 175-399, 175-461, 181-377, 181-391, 190-455, 192-455, 198-457, 200-409, 202-503, 203-458, 203-485, 204-536, 210-459, 211-410, 212-460, 218-342, 223-476, 227-430, 232-472, 234-442, 243-478, 244-459, 247-488, 249-455, 263-492, 267-468, 270-482, 270-483, 270-499, 272-400, 277-474, 285-454, 288-591, 290-548, 301-515, 303-499, 308-573
110/7506227CB1/ 795	1-215, 48-243, 118-249, 120-216, 120-219, 120-224, 120-249, 123-228, 123-249, 128-243, 128-249, 129-207, 132-239, 136-249, 142-249, 150-241, 151-437, 152-550, 159-249, 222-679, 246-441, 246-506, 248-522, 263-520, 285-508, 345-558, 345-605, 347-588, 360-603, 369-612, 379-630, 384-601, 386-600, 386-619, 388-617, 390-656, 402-679, 404-663, 411-621, 411-690, 434-667, 438-647, 445-692, 454-668, 454-691, 454-711, 473-642, 482-728, 482-734, 497-781, 501-732, 501-795, 506-730, 525-778, 642-778, 642-781
111/3144431CB1/ 895	1-431, 5-886, 92-744, 93-841, 110-272, 153-444, 166-434, 188-869, 235-484, 248-873, 438-872, 438-895, 498-799, 544-883, 617-895, 633-895
112/2633315CB1/ 2320	1-252, 1-849, 13-334, 27-639, 40-923, 49-856, 49-898, 51-753, 53-559, 60-622, 61-288, 61-697, 75-330, 75-869, 129-689, 129-714, 129-832, 249-952, 589-1146, 589-1213, 590-1182, 651-1367, 653-1310, 919-1381, 919-1456, 919-1485, 976-1192, 976-1520, 1043-1457, 1074-1479, 1168-1611, 1235-1480, 1235-1739, 1240-1435, 1290-1459, 1310-1766, 1318-1908, 1351-1603, 1351-1883, 1366-2127, 1367-1656, 1389-1983, 1430-2229, 1434-2048, 1446-1898, 1446-2141, 1472-1731, 1499-2000, 1503-1763, 1503-2005, 1530-2105, 1565-1782, 1631-1901, 1638-2204, 1678-2320, 1812-2131, 1822-2320
113/3401751CB1/ 1601	1-691, 1-1597, 470-886, 505-1142, 518-934, 682-1008, 720-807, 720-969, 720-1136, 720-1226, 720-1295, 720-1327, 721-958, 772-1088, 803-1345, 807-1258, 859-1169, 862-1103, 862-1164, 862-1373, 871-1312, 878-1216, 893-1460, 904-1571, 906-1435, 929-1601



Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
114/045680CB1/ 2857	1-188, 1-364, 37-125, 128-335, 128-692, 205-632, 223-365, 364-502, 397-874, 500-1068, 502-1329, 621-676, 672-874, 700-874, 871-1530, 1038-1383, 1090-1743, 1314-1724, 1336-1564, 1336-1646, 1336-1696, 1336-1724, 1336-1731, 1336-1788, 1336-1798, 1336-1824, 1336-1833, 1336-1848, 1336-1932, 1336-1940, 1336-1957, 1336-2010, 1336-2011, 1336-2045, 1337-1978, 1370-2015, 1428-2150, 1452-2187, 1453-2151, 1476-2153, 1477-2072, 1513-2237, 1545-1821, 1561-2260, 1568-2245, 1631-2163, 1642-2349, 1720-2086, 1769-2371, 1799-1977, 1799-2083, 1799-2170, 1799-2242, 1799-2306, 1799-2307, 1799-2534, 1811-2175, 1825-2091, 1931-2582, 1943-2211, 1949-2578, 1982-2653, 1994-2442, 1996-2526, 2030-2530, 2034-2535, 2036-2593, 2046-2565, 2097-2373, 2097-2391, 2104-2455, 2119-2781, 2159-2667, 2193-2500, 2210-2750, 2217-2835, 2241-2515, 2248-2424, 2305-2525, 2367-2852, 2408-2835, 2411-2668, 2411-2857, 2430-2857, 2431-2711, 2431-2855, 2453-2638, 2454-2855, 2459-2855, 2461-2821, 2485-2790, 2526-2847, 2529-2846, 2552-2857, 2577-2826, 2594-2855, 2596-2840, 2605-2854, 2651-2846, 2704-2840, 2793-2855
115/1503172CB1/ 1469	1-590, 4-816, 217-830, 266-776, 266-777, 290-816, 342-860, 345-945, 351-860, 374-867, 383-902, 422-1065, 469-1097, 498-910, 498-1110, 509-1021, 511-1140, 528-1053, 552-954, 621-1110, 629-1158, 675-1142, 682-1199, 694-1155, 694-1186, 694-1202, 700-1054, 721-1383, 731-1200, 763-1298, 824-1331, 871-1469, 951-1416, 1030-1200
116/1818665CB1/ 4976	1-720, 51-402, 173-383, 188-427, 204-422, 222-4976, 592-1161, 718-1285, 840-1403, 997-1666, 1056-1180, 1107-1697, 1174-1701, 1259-1701, 1331-2108, 1341-1941, 1341-2108, 1368-1975, 1573-2400, 1617-2161, 1726-2360, 1816-2013, 1816-2194, 1971-2445, 2031-2350, 2070-2660, 2201-2333, 2212-2673, 2245-2852, 2253-2621, 2304-2597, 2460-2603, 2509-2726, 2536-3176, 2631-3117, 2653-2919, 2659-3284, 2750-3570, 2777-3558, 2790-3383, 2806-3063, 2970-3595, 2980-3383, 2980-3611, 2980-3659, 2980-3681, 2992-3289, 3019-3313, 3019-3638, 3058-3559, 3063-3626, 3098-3861, 3099-3815, 3237-4451, 3243-3484, 3305-3448, 3307-3848, 3349-3596, 3349-3799, 3362-4012, 3400-3772, 3402-3671, 3455-3710, 3526-3921, 3528-4014, 3534-3994, 3550-4128, 3580-4211, 3598-4242, 3712-4005, 3728-4142, 3729-4307, 3734-4085, 3821-4452, 3883-4158, 3952-4451, 3952-4630, 4009-4513, 4098-4559, 4223-4719, 4230-4902, 4241-4771, 4251-4487, 4251-4680, 4251-4763, 4251-4806, 4252-4686, 4287-4921, 4291-4649, 4291-4881, 4295-4872, 4297-4897, 4304-4862, 4306-4549, 4322-4961, 4368-4877, 4380-4944, 4382-4955, 4385-4976, 4405-4941, 4412-4668, 4413-4976, 4443-4746, 4464-4894, 4467-4976, 4475-4720, 4475-4975, 4507-4976, 4513-4973, 4516-4969, 4518-4766, 4518-4968, 4518-4973, 4519-4962, 4523-4973, 4524-4969, 4525-4955, 4525-4969, 4530-4976, 4531-4976, 4532-4969, 4532-4976, 4534-4788, 4534-4956, 4534-4976, 4536-4976, 4537-4792, 4539-4974, 4553-4755, 4553-4962, 4553-4976, 4569-4969, 4582-4668, 4582-4970, 4588-4969, 4589-4969, 4594-4968, 4599-4969, 4608-4969, 4617-4924, 4669-4967, 4681-4958, 4681-4969, 4682-4959, 4700-4940, 4710-4834, 4735-4976, 4757-4973, 4761-4967, 4798-4969, 4828-4975, 4837-4917, 4859-4969

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
117/3251352CBI/ 3226	1-652, 1-683, 146-678, 338-880, 359-610, 359-636, 502-1371, 550-1100, 690-1115, 725-1045, 725-1262, 916-1111, 923-1214, 923-1566, 944-1152, 1118-1801, 1145-1395, 1145-1621, 1163-1914, 1205-1883, 1263-1476, 1297-1606, 1311-1401, 1343-2015, 1437-2009, 1440-2067, 1455-1926, 1477-1709, 1572-1831, 1589-2220, 1589-2242, 1589-2252, 1589-2277, 1621-1820, 1623-1907, 1623-2146, 1674-1902, 1712-1969, 1712-2252, 1730-1974, 1730-2328, 1738-2015, 1750-1907, 1800-2480, 1807-2427, 1836-2063, 1845-2495, 1881-2138, 1893-2349, 1910-2177, 1919-2332, 1948-2385, 1955-2256, 1977-2519, 1979-2503, 1991-2294, 2015-2493, 2033-2679, 2042-2306, 2049-2332, 2058-2462, 2080-2524, 2095-2530, 2098-2515, 2104-2344, 2104-2358, 2104-2412, 2104-2490, 2104-2511, 2104-2530, 2109-2761, 2174-2741, 2200-2462, 2203-2462, 2234-2462, 2246-2462, 2264-2528, 2265-2462, 2319-3022, 2341-2462, 2362-3100, 2403-3048, 2403-3226, 2406-2477, 2408-3089, 2417-2641, 2419-2677, 2508-3156, 2600-3086, 2773-3221
118/55091643CBI/ 1260	1-259, 1-514, 1-563, 1-575, 1-603, 1-688, 4-633, 28-598, 81-1260, 115-609, 162-587, 176-776, 223-468, 225-787, 225-818, 225-852, 225-855, 225-874, 227-836, 227-874, 230-864, 240-812, 245-839, 256-776, 268-864, 276-906, 280-848, 309-966, 384-632, 417-920, 419-1031, 447-962, 459-625, 481-994, 523-1197, 552-797, 596-1129
119/7500770CBI/ 649	1-154, 1-649, 18-205, 35-205, 134-626, 197-646, 197-649, 202-419, 202-429, 202-625, 202-628, 202-640, 202-642, 202-649, 203-423, 204-647, 207-649, 209-626, 211-648, 214-638, 220-636, 225-644, 240-642, 243-649, 244-643, 245-649, 246-640, 247-642, 248-649, 255-455, 268-642, 283-626, 283-637, 283-643, 283-644, 284-577, 288-647, 289-549, 290-642, 294-642, 294-645, 298-649, 303-644, 307-646, 313-642, 313-649, 348-638, 351-645, 369-642, 373-642, 378-633, 378-646, 378-649, 385-617, 385-642, 386-641, 390-572, 390-574, 392-577, 396-649, 397-649, 419-642, 424-649, 432-625, 441-626, 449-644, 461-641, 461-644, 461-649, 478-642, 483-647, 509-649, 541-645
120/7506350CBI/ 3439	1-282, 8-338, 8-3439, 23-273, 36-321, 39-275, 39-447, 39-509, 39-534, 39-561, 39-602, 39-605, 39-611, 39-625, 39-659, 39-661, 39-664, 39-673, 40-318, 40-355, 40-359, 47-341, 52-636, 61-544, 68-629, 70-661, 72-617, 72-630, 72-646, 72-942, 75-297, 75-336, 75-605, 80-640, 124-379, 237-817, 395-954, 421-714, 426-719, 444-716, 444-719, 456-760, 504-1069, 507-949, 540-1132, 565-1089, 602-997, 608-928, 610-814, 610-1276, 620-770, 621-884, 630-1269, 634-1088, 654-1087, 654-1121, 669-1136, 669-1208, 669-1300, 671-970, 671-1212, 676-1089, 717-1088, 733-1086, 737-1431, 784-1390, 786-1276, 872-1071, 872-1088, 1076-1661, 1083-1576, 1087-1349, 1087-1529, 1129-1530, 1130-1415, 1154-1707, 1248-1805, 1297-1527, 1369-1773, 1369-1805, 1397-1909, 1441-1595, 1455-1825, 1462-1710, 1482-1753, 1482-1763, 1537-2189, 1601-1805, 1797-1976, 1797-2082, 1797-2110, 1797-2417, 1847-2379, 1887-2214, 1916-2136, 1916-2318, 1927-2187, 1956-2195, 1966-2402, 1996-2822, 2012-2286, 2019-2560, 2038-2448, 2039-2453, 2058-2295, 2061-2446, 2063-2267, 2092-2563, 2099-2565, 2121-2565, 2141-2565, 2188-2561, 2193-2565, 2200-2514, 2260-3074, 2275-2537, 2275-2562, 2279-2865, 2351-2628, 2356-2616, 2356-2912, 2393-3021, 2394-2646, 2428-2905, 2430-2882, 2432-2830, 2449-2563, 2452-3084, 2464-3085, 2488-3054,

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
	2503-3146, 2550-2790, 2550-2808, 2550-2843, 2560-3060, 2577-3146, 2588-2895, 2588-3144, 2615-2882, 2634-3134, 2652-3144, 2652-3145, 2653-3146, 2654-3134, 2658-3130, 2660-3133, 2667-3135, 2671-3130, 2671-3146, 2673-3130, 2674-3131, 2678-3131, 2684-3131, 2685-3131, 2685-3134, 2687-3130, 2690-3132, 2690-3133, 2693-3136, 2696-3131, 2698-3131, 2704-3131, 2717-3131, 2724-3131, 2726-3131, 2731-3135, 2732-3131, 2735-3131, 2742-3028, 2744-3134, 2756-3294, 2765-3134, 2773-3134, 2775-3090, 2778-3270, 2781-3131, 2790-3027, 2809-3131, 2836-3131, 2872-3139, 3015-3439, 3171-3415
121/7508370CB1/ 1570	1-259, 1-310, 1-703, 1-711, 1-726, 1-750, 1-973, 278-994, 278-1021, 401-468, 668-1519, 1176-1570
122/2894093CB1/ 2671	1-533, 1-2654, 201-621, 243-618, 245-509, 245-713, 245-816, 245-833, 247-836, 248-508, 316-571, 602-1281, 619-2654, 691-808, 693-1295, 724-1169, 735-1352, 740-1319, 764-1335, 782-1330, 794-1350, 794-1404, 818-1390, 831-1475, 836-1413, 879-1521, 928-1620, 968-1379, 1066-1209, 1342-1927, 1355-1952, 1436-1656, 1458-1982, 1552-2064, 1570-2064, 1571-1756, 1770-2207, 1988-2262, 2185-2671
123/7507335CB1/ 3376	1-778, 637-1248, 637-1430, 643-935, 643-1054, 643-1178, 643-1233, 643-1306, 643-1359, 645-1335, 645-1389, 650-869, 650-937, 650-1151, 650-1334, 650-1466, 651-1269, 651-1303, 651-1431, 652-849, 652-1202, 652-1254, 654-894, 654-1029, 654-3356, 655-982, 656-1148, 656-1193, 657-924, 659-931, 659-1307, 659-1415, 660-910, 660-1288, 660-1302, 661-852, 661-889, 661-964, 661-1066, 661-1304, 661-1530, 662-1271, 663-1231, 663-1453, 664-979, 664-1308, 668-822, 668-1305, 668-1306, 668-1318, 668-1377, 670-942, 670-1302, 673-1443, 675-1280, 678-1305, 680-1291, 684-1228, 685-1115, 685-1118, 686-916, 687-1294, 694-1185, 700-1116, 730-1405, 730-1410, 754-1381, 767-1430, 856-1388, 861-1395, 876-1402, 905-1342, 905-1403, 905-1408, 972-1534, 1005-1506, 1052-1572, 1052-1596, 1052-1606, 1052-1608, 1086-1711, 1118-1636, 1119-1636, 1142-1636, 1157-1632, 1157-1636, 1179-1738, 1180-1711, 1180-1736, 1180-1738, 1181-1713, 1216-1719, 1233-1711, 1240-1681, 1252-1890, 1258-1894, 1261-1693, 1316-1711, 1316-1712, 1316-1738, 1343-1830, 1343-1883, 1343-1892, 1343-1909, 1344-1788, 1426-2156, 1466-2032, 1493-2107, 1522-2257, 1530-2151, 1534-2171, 1609-2107, 1609-2119, 1609-2127, 1609-2161, 1635-2107, 1642-2169, 1698-2510, 1749-2474, 1755-2355, 1755-2356, 1755-2406, 1764-2290,

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
	1776-2336, 1777-2322, 1779-2269, 1788-2467, 1793-2403, 1844-2418, 1858-2602, 1859-2371, 1864-2430, 1963-2503, 1987-2416, 1998-2640, 2007-2647, 2017-2566, 2018-2628, 2042-2611, 2055-2613, 2058-2614, 2066-2638, 2082-2629, 2129-2685, 2140-2802, 2140-2822, 2146-2625, 2146-2747, 2184-2736, 2218-2770, 2219-2700, 2246-2710, 2267-2787, 2275-2895, 2279-2772, 2338-2826, 2369-2932, 2485-2951, 2498-3123, 2512-3062, 2520-3051, 2520-3136, 2525-3210, 2553-3288, 2569-3052, 2599-3352, 2606-3223, 2660-3279, 2721-3179, 2732-3296, 2735-3375, 2737-3232, 2748-3356, 2759-3324, 2781-3218, 2797-3275, 2805-3352, 2820-3323, 2831-3356, 2837-3318, 2841-3301, 2845-3376, 2855-3301, 2861-3357, 2862-3356, 2874-3376, 2875-3335, 2875-3344, 2875-3356, 2876-3357, 2877-3376, 2878-3350, 2881-3374, 2887-3301, 2904-3356, 2913-3373, 2915-3356, 2917-3346, 2917-3352, 2917-3355, 2918-3352, 2925-3350, 2927-3352, 2932-3343, 2932-3350, 2935-3356, 2941-3356, 2942-3356
124/7509081CBI/ 2198	1-216, 1-255, 1-322, 1-416, 1-823, 12-643, 13-824, 23-777, 31-631, 33-164, 33-243, 33-415, 33-427, 33-560, 33-572, 33-596, 33-598, 33-613, 33-691, 33-784, 38-201, 39-672, 43-690, 45-290, 73-274, 73-700, 74-506, 98-479, 103-413, 134-405, 156-833, 186-817, 217-578, 226-483, 240-579, 259-963, 284-449, 306-814, 314-558, 314-803, 314-1003, 314-1020, 314-1021, 314-1025, 315-948, 317-947, 319-647, 322-491, 332-556, 335-843, 341-694, 349-744, 349-919, 350-902, 350-929, 365-1021, 368-605, 370-869, 371-826, 375-1062, 376-944, 380-679, 381-679, 395-1120, 396-811, 397-631, 399-881, 416-1025, 420-953, 421-701, 421-944, 424-1025, 429-684, 438-917, 442-950, 443-970, 456-1025, 460-1167, 470-973, 484-1104, 485-1105, 514-1188, 519-1099, 524-946, 528-1196, 537-1057, 537-1061, 538-1222, 544-1081, 544-1212, 554-872, 562-849, 563-1025, 583-920, 588-1100, 589-840, 591-1200, 596-918, 600-1222, 607-921, 612-1126, 624-952, 630-977, 635-1262, 655-1365, 662-962, 668-1025, 669-1239, 672-1224, 677-1321, 702-1300, 705-882, 705-1153, 705-1231, 708-995, 727-1045, 745-1407, 746-1203, 755-1394, 756-1178, 762-1363, 764-1379, 789-1129, 791-1233, 792-1034, 792-1042, 799-1407,
	816-1426, 837-1104, 837-1105, 849-1226, 854-1241, 865-1521, 875-1365, 932-1409, 932-1410, 943-1487, 981-1625, 1021-1420, 1021-1478, 1021-1720, 1021-1736, 1021-1798, 1026-1230, 1031-1763, 1031-1770, 1031-2198, 1032-1719, 1033-1573
125/7502450CBI/ 1453	1-470, 1-492, 1-552, 7-513, 16-568, 16-592, 21-689, 28-585, 32-620, 34-702, 47-696, 56-89, 63-707, 120-707, 160-707, 167-704, 169-456, 169-707, 170-697, 177-453, 192-707, 198-488, 201-707, 235-532, 258-532, 261-550, 261-557, 261-558, 261-561, 261-568, 263-563, 263-571, 264-935, 266-548, 270-549, 272-843, 272-907, 273-707, 278-707, 278-867, 283-647, 284-779, 286-695, 290-573, 290-704, 290-707, 291-707, 295-707, 296-586, 297-707, 298-833, 298-999, 299-951, 302-1074, 308-590, 311-1028, 313-707, 314-758, 317-609, 319-655, 320-704, 322-813, 322-935, 326-678, 328-678, 338-653, 338-707, 340-707, 341-652, 341-654, 344-707, 345-612, 350-966, 351-1011, 353-1080, 355-668, 357-630, 357-971, 357-1085, 358-707, 359-656, 360-667, 380-707, 387-657, 391-707, 394-707, 394-816, 415-707, 417-707, 418-975, 419-1024, 425-854, 428-707, 437-742, 439-707, 443-707, 447-745, 471-1069, 495-989, 495-1028, 514-813, 519-830, 555-1199, 578-1453, 596-1155, 616-898, 616-1438, 622-1278, 623-1185, 631-1187, 640-1232, 640-1239, 658-983, 661-1121, 664-983, 676-1216, 689-979, 690-980, 692-1173, 693-1343, 697-980, 700-1136, 703-898, 705-1038, 713-1443, 724-1036, 733-1326,

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
	739-1371, 755-1451, 763-1022, 763-1026, 773-1156, 773-1414, 783-1231, 785-1325, 786-1362, 837-1112, 844-1451, 849-1061, 851-1401, 862-1453, 863-1205, 863-1294, 863-1412, 863-1421, 863-1426, 863-1436, 863-1444, 863-1450, 863-1451, 863-1453, 864-1082, 864-1132, 869-1441, 907-1286, 913-1432, 927-1188, 984-1448, 987-1440, 992-1449, 993-1447, 1016-1443, 1017-1434, 1017-1453, 1022-1449, 1023-1291, 1025-1443, 1025-1453, 1026-1444, 1028-1446, 1035-1446, 1038-1449, 1058-1364, 1058-1369, 1058-1422, 1058-1453, 1065-1443, 1067-1443, 1068-1449, 1074-1445, 1076-1443, 1082-1430, 1103-1453, 1107-1153, 1107-1440, 1111-1453, 1112-1453, 1114-1443, 1119-1453, 1125-1446, 1128-1451, 1138-1431, 1143-1430, 1147-1446, 1150-1443, 1158-1453, 1179-1448
126/7501405CB1/ 949	1-567, 18-80, 24-636, 154-609, 168-781, 174-636, 178-513, 200-949, 230-636, 251-519, 302-615, 335-636, 387-558
127/7504528CB1/ 1265	1-332, 10-266, 10-361, 10-468, 10-470, 10-566, 10-587, 10-707, 10-720, 10-759, 10-780, 10-783, 10-796, 10-798, 10-799, 10-818, 10-819, 10-1257, 20-208, 20-282, 47-275, 64-351, 74-583, 95-653, 118-709, 146-545, 226-818, 226-900, 226-945, 228-522, 228-711, 228-781, 230-872, 249-474, 249-783, 249-818, 814-1261, 815-1265, 818-1123, 880-1150, 886-1192, 906-1172, 1049-1208
128/7509049CB1/ 703	1-248, 1-376, 1-413, 1-418, 1-433, 1-440, 1-461, 1-703, 5-239, 31-291, 32-299, 34-241, 34-292, 38-308, 38-626, 53-349, 58-414, 58-587, 63-328, 92-314, 144-444, 163-641, 164-441, 200-434, 211-435, 285-571, 299-423, 302-567, 302-579
129/7509086CB1/ 1165	1-1165, 20-469, 61-267, 70-452, 79-437, 80-155, 80-641, 84-469, 91-400, 107-469, 110-444, 111-196, 111-354, 111-429, 111-437, 111-465, 111-469, 117-710, 123-444, 125-431, 138-469, 196-455, 243-361, 274-409, 278-405, 308-406, 317-469, 432-909, 434-976, 444-976, 476-999, 528-1081, 553-1135, 600-949, 624-1164, 631-1116, 676-1118, 704-949, 705-1165, 706-941, 726-950, 740-1003, 742-1165, 749-1154, 754-1156, 755-1154, 756-1150, 780-1154, 785-1054, 792-1154, 808-1066, 810-1141, 811-1051, 815-1165, 817-1076, 817-1082, 819-1161, 822-1152, 832-1116, 855-1101, 855-1115, 859-1160, 860-939, 860-940, 860-947, 860-970, 860-1035, 860-1116, 860-1120, 861-883, 866-1042, 868-1095, 884-1154, 887-1153, 894-1154, 901-1165, 903-1154, 903-1163, 904-1164, 909-1155, 911-1154, 922-1150, 925-1165, 936-1153, 947-1154, 952-1150, 961-1155, 969-1149, 979-1114, 985-1162, 993-1163
130/7506914CB1/ 1625	1-276, 1-1625, 274-878, 526-1426, 565-1538, 569-1436, 597-1439, 604-1538, 610-1439, 617-1439, 618-1538, 620-1538, 623-1539, 674-1538, 683-977, 725-967, 729-1538, 739-1539, 741-1025, 745-1538, 761-1043, 772-1439, 778-989, 778-1539, 783-1539, 821-1538, 919-1538, 1093-1487, 1150-1408, 1274-1568, 1351-1625, 1353-1611, 1364-1618
131/5606114CB1/ 843	1-127, 1-226, 23-226, 31-224, 31-226, 33-226, 36-226, 37-226, 38-130, 38-158, 53-226, 54-226, 67-226, 67-245, 67-263, 67-305, 67-308, 67-349, 67-476, 71-441, 72-341, 73-346, 86-830, 99-318, 155-430, 301-755, 344-783, 373-766, 395-784, 403-783, 404-783, 429-796, 432-783, 432-784, 437-792, 444-786, 445-783, 446-784, 447-783, 448-783, 460-783, 466-843, 467-694, 467-843, 469-783, 479-781, 481-620, 481-716, 512-783, 548-783, 581-798, 590-783, 599-783, 609-784, 707-843, 725-793

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
132/7503282CB1/ 1589	1-244, 1-246, 1-487, 1-1589, 2-143, 26-192, 38-487, 53-293, 54-297, 64-487, 104-487, 284-429, 298-975, 399-1126, 486-778, 487-648, 487-779, 487-913, 487-917, 487-951, 487-1001, 487-1088, 487-1100, 495-1062, 504-777, 511-1090, 511-1140, 518-844, 519-1085, 522-1123, 526-1095, 535-1101, 536-1083, 538-1050, 549-895, 550-911, 555-1162, 563-975, 569-1008, 572-1065, 578-900, 578-925, 582-1411, 587-1017, 594-981, 594-1047, 594-1070, 600-1070, 605-1048, 606-1077, 611-1120, 622-938, 633-1291, 644-1077, 664-943, 665-1005, 665-1186, 669-1240, 671-1422, 672-1197, 675-944, 685-1122, 698-1386, 702-1065, 711-1445, 719-982, 720-1123, 721-1018, 726-1340, 729-890, 732-1337, 733-1027, 733-1174, 734-1015, 735-1295, 759-1348, 761-1325, 770-1404, 773-1186, 787-1070, 798-1221, 801-1020, 803-1407, 807-1311, 809-1130, 817-1306, 819-1435, 831-1028, 838-1436, 844-1431, 847-1315, 847-1348, 847-1368, 855-1437, 873-1445, 880-1309, 880-1348, 883-1348, 886-1367, 907-1169, 912-1474, 912-1563, 912-1564, 912-1588, 913-1484, 914-1046, 918-1044, 918-1321, 918-1348, 922-1164, 925-1417, 925-1574, 954-1245, 969-1414, 972-1445, 976-1253, 977-1381, 977-1427, 989-1445, 991-1446,
	995-1447, 996-1442, 1004-1445, 1011-1445, 1014-1272, 1018-1270, 1020-1444, 1022-1444, 1025-1475, 1033-1437, 1035-1444, 1049-1445, 1050-1445, 1057-1331, 1064-1445, 1067-1445, 1075-1445, 1078-1348, 1083-1444, 1085-1413, 1086-1444, 1087-1444, 1089-1349, 1089-1444, 1090-1445, 1102-1444, 1102-1445, 1110-1393, 1112-1442, 1122-1445, 1155-1379, 1155-1444, 1177-1446, 1209-1445, 1254-1445, 1329-1459, 1331-1445, 1333-1445, 1334-1431, 1338-1459, 1339-1444, 1341-1589, 1342-1445, 1347-1442
133/7503284CB1/ 1748	1-244, 1-246, 1-487, 1-1748, 2-143, 26-192, 38-487, 53-293, 54-297, 64-487, 104-487, 284-429, 314-1024, 336-790, 336-922, 407-876, 407-887, 409-886, 434-676, 436-961, 450-1005, 469-1141, 487-696, 487-717, 487-802, 487-911, 487-967, 497-720, 502-1120, 552-1096, 572-1132, 580-1038, 582-1160, 585-845, 585-1058, 585-1076, 585-1101, 585-1212, 585-1308, 587-1135, 595-1160, 602-1155, 602-1169, 603-796, 609-1132, 621-1072, 624-807, 624-1110, 626-1076, 633-938, 634-1160, 646-937, 663-936, 677-1003, 708-1054, 709-1070, 722-1134, 737-1084, 781-1097, 824-1164, 878-1141, 1072-1643, 1173-1431, 1488-1618, 1497-1618, 1500-1748
134/7510501CB1/ 621	1-438, 42-602, 58-162, 66-162, 71-162, 72-162, 73-162, 88-162, 89-162, 102-162, 103-325, 106-162, 107-162, 108-162, 109-162, 116-162, 126-162, 141-620, 163-222, 163-240, 163-241, 163-260, 163-279, 163-285, 163-339, 163-411, 163-424, 163-439, 163-449, 163-470, 163-546, 163-585, 163-590, 163-602, 163-612, 165-602, 166-602, 167-602, 177-609, 183-602, 192-585, 200-612, 214-602, 237-464, 248-602, 251-602, 256-602, 285-602, 288-602, 298-607, 380-621, 400-602, 409-607, 428-613

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
135/7500444CB1/ 2220	1-823, 1-2220, 108-516, 108-530, 108-777, 154-426, 154-837, 154-850, 156-411, 177-771, 191-623, 195-828, 213-785, 216-718, 244-389, 244-579, 251-442, 251-638, 251-805, 264-757, 344-499, 384-1128, 390-819, 394-901, 406-662, 420-736, 420-871, 431-699, 432-746, 433-828, 468-717, 478-780, 493-1096, 500-803, 509-1110, 534-754, 534-1116, 537-815, 551-810, 567-826, 568-1134, 572-847, 582-1078, 584-1048, 590-961, 590-995, 590-1043, 596-839, 600-1037, 603-1083, 619-812, 646-1134, 647-1148, 684-898, 693-927, 693-960, 693-965, 709-1136, 724-947, 724-993, 724-1124, 740-1136, 746-883, 748-1008, 781-1123, 786-983, 786-1035, 803-1074, 812-1045, 813-1101, 813-1377, 818-1057, 819-1022, 819-1057, 820-1088, 835-1037, 862-1075, 864-1121, 864-1136, 913-1480, 946-1690, 977-1134, 1081-1300, 1093-1433, 1371-1534, 1374-1836, 1376-2081, 1407-1649, 1421-1841, 1438-1704, 1480-1840, 1550-1824, 1583-1853, 1639-1840, 1811-2055, 1818-1921, 1865-2210
136/7510297CB1/ 1990	1-142, 1-1985, 88-377, 89-308, 89-318, 89-370, 89-427, 89-445, 89-492, 89-506, 89-508, 89-519, 89-533, 89-542, 89-626, 89-634, 91-142, 91-358, 91-425, 91-546, 93-551, 94-626, 102-368, 113-556, 114-360, 124-582, 129-595, 139-675, 179-692, 191-514, 198-715, 203-793, 218-475, 232-390, 235-758, 242-714, 251-791, 281-722, 284-559, 287-549, 312-594, 322-600, 322-707, 322-878, 322-885, 322-1011, 324-592, 327-621, 328-850, 364-671, 364-847, 392-923, 395-644, 397-907, 397-908, 407-1086, 411-738, 413-928, 423-660, 423-667, 423-876, 432-682, 467-973, 483-973, 484-867, 484-1259, 485-779, 485-954, 485-1013, 485-1054, 485-1061, 485-1074, 485-1152, 485-1183, 485-1206, 485-1222, 487-1189, 497-763, 507-1063, 511-1063, 523-773, 537-755, 537-813, 560-820, 564-1068, 578-900, 579-1112, 583-1030, 584-936, 591-1153, 596-858, 598-1101, 601-874, 601-1284, 607-923, 607-1086, 607-1224, 609-1071, 613-1450, 621-770, 621-950, 623-1026, 629-879, 632-911, 637-1161, 640-1146, 642-1172, 659-945, 659-1131, 661-1467, 673-1199, 674-1161, 674-1165, 675-900, 693-801, 694-801, 711-1208, 725-971, 725-988, 725-1197, 725-1221, 725-1222, 725-1250, 725-1259, 725-1281, 725-1334, 729-1362, 731-784, 733-1220, 735-1231, 747-1450, 761-1269, 763-1029, 765-1286, 768-1268, 770-1036, 773-1063, 774-1260, 774-1312, 791-1022, 792-1378, 802-1368, 808-1309, 808-1326, 809-1075, 811-1092, 812-1367, 815-1014, 815-1032, 815-1092, 824-1028, 825-1563, 829-1344, 830-1070, 831-1099, 837-1074, 841-1398, 845-1127, 846-1028, 858-1324, 859-1373, 860-1446, 860-1450, 863-1074, 865-1118, 865-1325, 865-1326, 865-1340, 865-1353, 865-1358, 865-1450, 865-1519, 865-1552, 866-1450, 869-1496, 871-1359, 874-1055, 877-1685, 882-1343, 897-1614, 898-1481, 908-1123, 922-1191, 928-1252, 940-1547, 942-1489, 947-1158, 947-1441, 948-1187, 960-1195, 961-1466, 965-1589, 980-1449, 982-1497, 983-1343, 983-1505, 994-1361, 1001-1267, 1001-1268, 1001-1269, 1002-1450, 1004-1501, 1009-1288, 1009-1450, 1010-1589, 1013-1450, 1014-1449, 1016-1114, 1016-1244, 1016-1463, 1019-1443, 1025-1450, 1025-1731, 1030-1340, 1030-1654, 1030-1819, 1031-1279, 1031-1287, 1034-1329, 1037-1438, 1037-1601, 1037-1629, 1037-1642, 1037-1723, 1037-1842, 1041-1842, 1043-1575, 1043-1594, 1045-1321, 1047-1586, 1048-1579, 1052-1553, 1052-1614, 1057-1671, 1058-1406, 1058-1560, 1058-1643, 1061-1631, 1062-1450, 10

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
	1063-1656, 1063-1751, 1064-1340, 1065-1679, 1066-1394, 1081-1576, 1086-1741, 1092-1638, 1105-1322, 1105-1323, 1107-1840, 1109-1840, 1110-1588, 1111-1321, 1111-1843, 1113-1330, 1114-1382, 1114-1391, 1119-1452, 1119-1839, 1125-1702, 1126-1456, 1128-1726, 1129-1382, 1129-1631, 1129-1729, 1133-1731, 1141-1393, 1144-1434, 1144-1985, 1146-1450, 1146-1489, 1152-1798, 1154-1648, 1157-1432, 1161-1423, 1161-1749, 1162-1311, 1162-1450, 1162-1759, 1164-1450, 1164-1473, 1168-1450, 1172-1398, 1172-1441, 1173-1382, 1173-1842, 1175-1445, 1177-1719, 1178-1724, 1178-1980, 1180-1464, 1182-1823, 1185-1581, 1187-1985, 1189-1700, 1193-1762, 1195-1985, 1196-1709, 1198-1985, 1199-1783, 1200-1462, 1201-1735, 1209-1843, 1211-1684, 1214-1450, 1224-1428, 1224-1461, 1229-1515, 1229-1516, 1232-1604, 1232-1990, 1235-1985, 1238-1513, 1238-1537, 1240-1840, 1241-1798, 1242-1985, 1245-1453, 1248-1425, 1251-1985, 1253-1544, 1253-1836, 1255-1801, 1268-1551, 1269-1629, 1269-1794, 1272-1723, 1278-1826, 1279-1703, 1281-1985, 1284-1834, 1285-1574, 1286-1557, 1287-1800, 1287-1804, 1290-1844, 1291-1529, 1294-1551, 1295-1805, 1296-1837, 1297-1843, 1299-1513, 1299-1801, 1302-1548,
	1302-1811, 1306-1674, 1310-1806, 1310-1840, 1311-1579, 1311-1844, 1315-1844, 1324-1706, 1327-1723, 1331-1607, 1335-1843, 1338-1840, 1343-1813, 1344-1737, 1344-1745, 1344-1825, 1345-1798, 1347-1842, 1348-1588, 1348-1623, 1350-1842, 1357-1636, 1357-1844, 1358-1927, 1360-1723, 1360-1813, 1362-1523, 1363-1842, 1366-1839, 1367-1802, 1371-1653, 1371-1843, 1376-1583, 1379-1843, 1385-1549, 1385-1843, 1388-1826, 1390-1840, 1391-1709, 1391-1844, 1392-1815, 1392-1825, 1392-1981, 1393-1700, 1393-1839, 1394-1841, 1394-1843, 1396-1519, 1396-1642, 1397-1985, 1398-1843, 1399-1842, 1400-1841, 1400-1842, 1401-1839, 1401-1843, 1402-1553, 1402-1842, 1404-1840, 1404-1843, 1405-1538, 1406-1609, 1406-1638, 1406-1842, 1407-1843, 1407-1984, 1408-1839, 1410-1843, 1410-1844, 1411-1841, 1411-1843, 1418-1624, 1419-1840, 1421-1841, 1421-1843, 1423-1843, 1425-1949, 1427-1559, 1427-1841, 1428-1843, 1428-1984, 1429-1701, 1429-1843, 1430-1841, 1430-1843, 1431-1843, 1432-1843, 1439-1984, 1444-1828, 1444-1942, 1445-1716, 1445-1784, 1454-1843, 1457-1843, 1462-1984, 1465-1843, 1468-1843, 1469-1605, 1469-1843, 1470-1748, 1470-1826, 1470-1830, 1470-1840, 1470-1841, 1471-1710, 1473-1843,
	1475-1716, 1475-1736, 1477-1841, 1477-1843, 1478-1772, 1481-1843, 1488-1812, 1489-1843, 1492-1843, 1496-1984, 1499-1787, 1499-1840, 1499-1984, 1500-1843, 1502-1743, 1503-1843, 1505-1789, 1505-1841, 1506-1841, 1507-1843, 1509-1909, 1511-1842, 1519-1726, 1519-1889, 1519-1984, 1521-1690, 1521-1984, 1522-1843, 1523-1985, 1524-1841, 1527-1842, 1528-1944, 1530-1841, 1533-1984, 1536-1799, 1536-1842, 1537-1807, 1538-1835, 1538-1842, 1538-1864, 1539-1779, 1542-1984, 1546-1735, 1546-1843, 1553-1841, 1557-1817, 1559-1841, 1561-1817, 1561-1842, 1563-1842, 1563-1984, 1578-1984, 1582-1653, 1582-1939, 1583-1803, 1588-1842, 1590-1843, 1599-1842, 1599-1843, 1613-1841, 1614-1863, 1618-1843, 1621-1841, 1628-1984, 1631-1842, 1644-1841, 1652-1840, 1652-1843, 1655-1843, 1658-1841, 1659-1984, 1683-1840, 1690-1984, 1692-1984, 1699-1983, 1702-1841, 1702-1984, 1707-1984, 1709-1808, 1713-1842, 1714-1984, 1719-1985, 1721-1984, 1734-1984, 1747-1981, 1760-1984, 1774-1841, 1798-1947, 1798-1984, 1869-1990



Table 4

Polynucleotide SEQ ID NO.:/ Incyte ID/ Sequence Length	Sequence Fragments
137/7640560CB1/ 2164	1-2164, 197-885, 201-917, 201-2164, 213-992, 215-916, 229-926, 231-1066, 234-811, 301-854, 406-1007, 475-1093, 660-1232, 671-1060, 671-1198, 747-1570, 774-1403, 927-1488, 944-1516, 949-1419, 970-1584, 1043-1636, 1055-1654, 1128-1564, 1158-1649, 1189-1642, 1196-1656, 1215-1652, 1274-1647, 1379-1650, 1534-2164, 1549-1833
138/7506087CB1/ 1438	1-147, 1-203, 1-1239, 1-1397, 220-483, 220-812, 285-912, 315-945, 317-940, 345-925, 356-714, 359-714, 361-472, 386-1106, 391-817, 391-966, 398-571, 422-1059, 438-983, 439-708, 439-861, 440-993, 465-690, 472-1011, 508-1054, 508-1117, 520-1155, 567-1068, 589-1062, 601-1212, 613-1120, 624-1215, 675-1029, 698-1162, 722-1006, 729-872, 814-1213, 843-1043, 1005-1239, 1240-1438

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID:	Representative Library
70	1393336CB1	THYRNOT03
71	1431502CB1	BEPINON01
72	2445220CB1	FIBRTXS07
73	5504385CB1	PITUNON01
74	6974948CB1	BRABDIR01
75	7501636CB1	LIVRTUE01
76	2535717CB1	PLACNOT02
77	6119548CB1	SINTNOR01
78	72263451CB1	LUNPTMK03
79	7502640CB1	BRAGNON02
80	7505807CB1	BRABDIK02
81	7506413CB1	SINTNOR01
82	1283631CB1	KERANOT02
83	1740413CB1	GBLADIE01
84	1951731CB1	BRATNOR01
85	3741930CB1	MENTNOT01
86	5402506CB1	CONUTUT01
87	71081333CB1	BEPINON01
88	7503139CB1	TONGTUT01
89	7505836CB1	LUNGNON07
90	7505858CB1	PROSTUT05
91	7505872CB1	BRAYDIN03
92	7506456CB1	BRAUTDR04
93	7506697CB1	GBLADIE01
94	7623472CB1	SMCCNON03
95	7506416CB1	SINTNOR01
96	4823849CB1	PROSTUT17
97	4433922CB1	EPIPON05
98	7504597CB1	BRAHNON05
99	7505987CB1	COLNUCT03
100	7506025CB1	BRSTNOT01
101	7506102CB1	TYMNOT08
102	1333949CB1	COLNNOT13
103	7035533CB1	ADRETUT06
104	2815375CB1	BRAYDIN03
105	2820152CB1	LUNGFET03
106	2959305CB1	MCLDTXN03
107	4913449CB1	THYMDIT01
108	7506136CB1	UTRSTMC01
109	7506225CB1	TYMNOT08
110	7506227CB1	COLNFET02
111	3144431CB1	HNT2AZS07
112	2633315CB1	BMARTXE01
113	3401751CB1	MUSCDMT01
114	045680CB1	LUNGNON07
115	1503172CB1	BRAITUT07
116	1818665CB1	PITUNON01
117	3251352CB1	ADRENOT03
118	55091643CB1	TNFRDNV01

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID:	Representative Library
119	7500770CB1	PROSBPS05
120	7506350CB1	BRAYDIN03
121	7508370CB1	PROSNOT20
122	2894093CB1	KIDNTUT14
123	7507335CB1	SINITMC01
124	7509081CB1	BRSTTUT08
125	7502450CB1	FIBRTXS07
126	7501405CB1	COLNNOT09
127	7504528CB1	LIVRTUT04
128	7509049CB1	PENITUT01
129	7509086CB1	HEARNON03
130	7506914CB1	COLDNOT01
131	5606114CB1	ADRETUT06
132	7503282CB1	BRSTTUT01
133	7503284CB1	BRAINOT14
134	7510501CB1	UTREDMT07
135	7500444CB1	LUNGFET03
136	7510297CB1	NGANNOT01
137	7640560CB1	BRAIFER06
138	7506087CB1	OVARNOT07

Table 6

Library	Vector	Library Description
ADRENOT03	PSPORT1	Library was constructed using RNA isolated from the adrenal tissue of a 17-year-old Caucasian male, who died from cerebral anoxia.
ADRETUT06	pINCY	Library was constructed using RNA isolated from adrenal tumor tissue removed from a 57-year-old Caucasian female during a unilateral right adrenalectomy. Pathology indicated pheochromocytoma, forming a nodular mass completely replacing the medulla of the adrenal gland.
BEPINON01	pTTT3	Normalized library was constructed from 5.12 million independent clones from a bronchial epithelium library. RNA was made from a bronchial epithelium primary cell line derived from a 54-year-old Caucasian male. The normalization and hybridization conditions were adapted from Soares et al., PNAS (1994) 91:9228, using a longer (24-hour) reannealing hybridization period.
BMARTXE01	pINCY	This 5' biased random primed library was constructed using RNA isolated from treated SH-SY5Y cells derived from a metastatic bone marrow neuroblastoma, removed from a 4-year-old Caucasian female (Schering AG). The medium was MEM/HAM'S F12 with 10% fetal calf serum. After reaching about 80% confluency cells were treated with 6-Hydroxydopamine (6-OHDA) at 100 microM for 8 hours.
BRABDIK02	PSPORT1	This amplified and normalized library was constructed using pooled cDNA from three different donors. cDNA was generated using mRNA isolated from diseased vermis tissue removed from a 79-year-old Caucasian female (donor A) who died from pneumonia, an 83-year-old Caucasian male (donor B) who died from congestive heart failure, and an 87-year-old Caucasian female (donor C) who died from esophageal cancer. Pathology indicated severe Alzheimer's disease in donors A & B and moderate Alzheimer's disease in donor C. Patient history included glaucoma, pseudophakia, gastritis with gastrointestinal bleeding, peripheral vascular disease, chronic obstructive pulmonary disease, seizures, tobacco abuse in remission, and transitory ischemic attacks in donor A; Parkinson's disease and atherosclerosis in donor B; hypertension, coronary artery disease, cerebral vascular accident, and hypothyroidism in donor C. Family history included Alzheimer's disease in the mother and sibling(s) of donor A. Independent clones from this amplified library were normalized in one round using conditions adapted Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
BRABDIR01	pINCY	Library was constructed using RNA isolated from diseased cerebellum tissue removed from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident. Patient history included Huntington's disease, emphysema, and tobacco abuse.

Table 6

Library	Vector	Library Description
BRAGNON02	pINCY	This normalized substantia nigra tissue library was constructed from 4.2 10e7 independent clones from a substantia nigra tissue library. Starting RNA was made from substantia nigra tissue removed from an 81-year-old Caucasian female who died from a hemorrhage and ruptured thoracic aorta due to atherosclerosis. Pathology indicated moderate atherosclerosis involving the internal carotids, bilaterally; microscopic infarcts of the frontal cortex and hippocampus; and scattered diffuse amyloid plaques and neurofibrillary tangles, consistent with age. Grossly, the leptomeninges showed only mild thickening and hyalinization along the superior sagittal sinus. The remainder of the leptomeninges was thin and contained some congested blood vessels. Mild atrophy was found mostly in the frontal poles and lobes, and temporal lobes, bilaterally. Microscopically, there were pairs of Alzheimer type II astrocytes within the deep layers of the neocortex. There was increased satellitosis around neurons in the deep gray matter in the middle frontal cortex. The amygdala contained rare diffuse plaques and neurofibrillary tangles. The posterior hippocampus contained a microscopic area of cystic cavitation with hemosiderin-laden macrophages surrounded by reactive gliosis. Patient history included sepsis, cholangitis, post-operative atelectasis, pneumonia CAD, cardiomegaly due to left ventricular hypertrophy, splenomegaly, arteriolonephrosclerosis, nodular colloidal goiter, emphysema, CHF, hypothyroidism, and peripheral vascular disease. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
BRAHNON05	pINCY	This normalized hippocampus tissue library was constructed from 1.6 million independent clones from a hippocampus tissue library. Starting RNA was made from posterior hippocampus removed from a 35-year-old Caucasian male who died from cardiac failure. Pathology indicated moderate leptomeningeal fibrosis and multiple microinfarctions of the cerebral neocortex. The cerebral hemisphere revealed moderate fibrosis of the leptomeninges with focal calcifications. There was evidence of shrunken and slightly eosinophilic pyramidal neurons throughout the cerebral hemispheres. There were small microscopic areas of cavitation with gliosis, scattered through the cerebral cortex. Patient history included cardiomyopathy, CHF, cardiomegaly, an enlarged spleen and liver. Patient medications included simethicone, Lasix, Digoxin, Colace, Zantac, captopril, and Vasotec. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
BRAIFER06	PCDNA2.1	This random primed library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus who was stillborn with a hypoplastic left heart at 23 weeks' gestation. Serologies were negative.

Table 6

Library	Vector	Library Description
BRAINOT14	pINCY	Library was constructed using RNA isolated from brain tissue removed from the left frontal lobe of a 40-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology for the associated tumor tissue indicated grade 4 gemistocytic astrocytoma.
BRAITUT07	pINCY	Library was constructed using RNA isolated from left frontal lobe tumor tissue removed from the brain of a 32-year-old Caucasian male during excision of a cerebral meningeal lesion. Pathology indicated low grade desmoplastic neuronal neoplasm, type not otherwise specified. The lesion formed a firm, circumscribed cyst-associated mass involving white matter and cortex. No definite glial component was evident to suggest a diagnosis of ganglioglioma. Family history included atherosclerotic coronary artery disease.
BRATNOR01	PCDNA2.1	This random primed library was constructed using RNA isolated from temporal cortex tissue removed from a 45-year-old Caucasian female who died from a dissecting aortic aneurysm and ischemic bowel disease. Pathology indicated mild arteriosclerosis involving the cerebral cortical white matter and basal ganglia. Grossly, there was mild meningeal fibrosis and mild focal atherosclerotic plaque in the middle cerebral artery, as well as vertebral arteries bilaterally. Microscopically, the cerebral hemispheres, brain stem and cerebellum reveal focal areas in the white matter that contain blood vessels that were barrel-shaped, hyalinized, with hemosiderin-laden macrophages in the Virchow-Robin space. In addition, there were scattered neurofibrillary tangles within the basolateral nuclei of the amygdala. Patient history included mild atheromatosis of aorta and coronary arteries, bowel and liver infarct due to aneurysm, physiologic fatty liver associated with obesity, mild diffuse emphysema, thrombosis of mesenteric and portal veins, cardiomegaly due to hypertrophy of left ventricle, arterial hypertension, acute pulmonary edema, splenomegaly, obesity (300 lb.), leiomyoma of uterus, sleep apnea, and iron deficiency anemia.

Table 6

Library	Vector	Library Description
BRAUTDR04	PCDNA2.1	Library was constructed using 1.5 micrograms of polyA RNA isolated from striatum, dorsal caudate nucleus, dorsal putamen, and ventral nucleus accumbens tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated no diagnostic abnormalities in the brain or intracranial vessels. There was mild meningeal fibrosis predominately over the convexities. Special stains showed no evidence of amyloid plaques or metastatic lesions. There were scattered axonal spheroids in the white matter of the cingulate cortex and thalamus. There were a few scattered neurofibrillary tangles in the entorhinal cortex and periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor, surrounded by foci of bile lakes beneath the hepatic surface scar. The liver had extensive surface scarring, congestion, cholestasis, hemorrhage, necrosis, and chronic inflammation. The patient presented with nausea, vomiting, dehydration, malnutrition, oliguria, and acute renal failure. Patient history included post-operative Budd-Chiari syndrome, biliary ascites, acute bilateral bronchopneumonia with microabscesses, hydrothorax, and bilateral leg pitting edema. Previous surgeries included cholecystectomy, liver resection, hysterectomy, bilateral salpingo-oophorectomy, and portocaval shunt. The patient was treated with a nasogastric feeding tube, biliary drainage stent, paracentesis, pleurodesis, and abdominal ultrasound. Patient medications included Ampicillin, niacin, furosemide, Aldactone, Benadryl, and morphine.
BRAYDIN03	pINCY	This normalized library was constructed from 6.7 million independent clones from a brain tissue library. Starting RNA was made from RNA isolated from diseased hypothalamus tissue removed from a 57-year-old Caucasian male who died from a cerebrovascular accident. Patient history included Huntington's disease and emphysema. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 -hours/round) reannealing hybridization was used. The library was linearized and recircularized to select for insert containing clones.
BRSTNOT01	PBLUESCRIPT	Library was constructed using RNA isolated from the breast tissue of a 56-year-old Caucasian female who died in a motor vehicle accident.

Table 6

Library	Vector	Library Description
BRSTTUT01	PSPORT1	Library was constructed using RNA isolated from breast tumor tissue removed from a 55-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated invasive grade 4 mammary adenocarcinoma of mixed lobular and ductal type, extensively involving the left breast. The tumor was identified in the deep dermis near the lactiferous ducts with extracapsular extension. Seven mid and low and five high axillary lymph nodes were positive for tumor. Proliferative fibrocytic changes were characterized by apocrine metaplasia, sclerosing adenosis, cyst formation, and ductal hyperplasia without atypia. Patient history included atrial tachycardia, blood in the stool, and a benign breast neoplasm. Family history included benign hypertension, atherosclerotic coronary artery disease, cerebrovascular disease, and depressive disorder.
BRSTTUT08	pINCY	Library was constructed using RNA isolated from breast tumor tissue removed from a 45-year-old Caucasian female during unilateral extended simple mastectomy. Pathology indicated invasive nuclear grade 2-3 adenocarcinoma, ductal type, with 3 of 23 lymph nodes positive for metastatic disease. Greater than 50% of the tumor volume was in situ, both comedo and non-comedo types. Immunostains were positive for estrogen/progesterone receptors, and uninvolved tissue showed proliferative changes. The patient concurrently underwent a total abdominal hysterectomy. Patient history included valvuloplasty of mitral valve without replacement, rheumatic mitral insufficiency, and rheumatic heart disease. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
COLDNOT01	pINCY	Library was constructed using RNA isolated from diseased descending colon tissue removed from a 16-year-old Caucasian male during partial colectomy, temporary ileostomy, and colonoscopy. Pathology indicated innumerable (greater than 100) adenomatous polyps with low grade dysplasia involving the entire colonic mucosa in the setting of familial polyposis coli. The patient presented with abdominal pain and flatulence. The patient was not taking any medications. Family history included benign colon neoplasm in the father; benign colon neoplasm in the sibling(s); and benign hypertension, cerebrovascular disease, breast cancer, uterine cancer, and type II diabetes in the grandparent(s).
COLNFET02	pINCY	Library was constructed using RNA isolated from the colon tissue of a Caucasian female fetus, who died at 20 weeks' gestation.
COLNNOT09	PSPORT1	Library was constructed using RNA isolated from colon tissue removed from a 60-year-old Caucasian male during a left hemicolectomy.
COLNNOT13	pINCY	Library was constructed using RNA isolated from ascending colon tissue of a 28-year-old Caucasian male with moderate chronic ulcerative colitis.



Table 6

Library	Vector	Library Description
COLNUCT03	pINCY	Library was constructed using RNA isolated from diseased colon tissue obtained from a 69-year-old Caucasian male during a partial colon excision with ileostomy. Pathology indicated severely active idiopathic inflammatory bowel disease most consistent with chronic ulcerative colitis. Patient history included benign neoplasm of the colon. Previous surgeries included cholecystectomy, spinal canal exploration, partial glossectomy, radical cystectomy, and bladder operation. Family history included cerebrovascular disease and benign hypertension.
CONUTUT01	pINCY	Library was constructed using RNA isolated from sigmoid mesentery tumor tissue obtained from a 61-year-old female during a total abdominal hysterectomy and bilateral salpingo-oophorectomy with regional lymph node excision. Pathology indicated a metastatic grade 4 malignant mixed müllerian tumor present in the sigmoid mesentery at two sites.
EIPN05	pINCY	This normalized prostate epithelial cell tissue library was constructed from 2.36 million independent clones from a prostate epithelial cell tissue library. Starting RNA was made from untreated prostatic epithelial cell issue removed from a 17-year-old Hispanic male. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 -hours/round) reannealing hybridization was used.
FIBRTXS07	pINCY	This subtracted library was constructed using 1.3 million clones from a dermal fibroblast library and was subjected to two rounds of subtraction hybridization with 2.8 million clones from an untreated dermal fibroblast tissue library. The starting library for subtraction was constructed using RNA isolated from treated dermal fibroblast tissue removed from the breast of a 31-year-old Caucasian female. The cells were treated with 9CIS retinoic acid. The hybridization probe for subtraction was derived from a similarly constructed library from RNA isolated from untreated dermal fibroblast tissue from the same donor. Subtractive hybridization conditions were based on the methodologies of Swaroop et al., NAR (1991) 19:1954 and Bonaldo, et al., Genome Research (1996) 6:791.
GBLADIE01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from diseased gallbladder tissue removed from a 55-year-old Caucasian female during laparoscopic cholecystectomy. Pathology indicated chronic cholecystitis and cholelithiasis (greater than 100 stones). The patient presented with cholelithiasis, abdominal pain, and tremors. Patient history included benign hypertension, Morton's neuroma, facial hirsutism, normal delivery, and tobacco abuse in remission. Previous surgeries included total abdominal hysterectomy, bilateral salpingo-oophorectomy, and adenotonsillectomy. Patient medications included Inderal and Premarin. Family history included breast cancer and ALS in the mother; chronic leukemia and ARDS in the father; breast cancer in the sibling(s); and atherosclerotic coronary artery disease in the grandparent(s).

Table 6

Library	Vector	Library Description
HEARNON03	pINCY	This normalized heart tissue library was constructed from 8.4 million independent clones from a heart tissue library. Starting RNA was made from heart tissue removed from a 44-year-old Caucasian male, who died from intracranial hemorrhage. Serology was positive for anti-CMV (cytomegalovirus). Patient history included back and neck pain, hypertension, pneumonia, sinus infection, alcohol use, and daily pipe tobacco use (x3 years). Patient medications included Procridia. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
HNT2AZS07	PSPORT1	This subtracted library was constructed from RNA isolated from an hNT2 cell line (derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor) treated for three days with 0.35 micromolar AZ. The hybridization probe for subtraction was derived from a similarly constructed library from untreated hNT2 cells. 3.08M clones from the AZ-treated library were subjected to three rounds of subtractive hybridization with 3.04M clones from the untreated library. Subtractive hybridization conditions were based on the methodologies of Swaroop et al. (NAR (1991) 19:1954) and Bonaldo et al. (Genome Research (1996) 6:791).
KERANOT02	PSPORT1	Library was constructed using RNA isolated from epidermal breast keratinocytes (NHEK). NHEK (Clontech #CC-2501) is human breast keratinocyte cell line derived from a 30-year-old black female during breast-reduction surgery.
KIDNTUT14	pINCY	Library was constructed using RNA isolated from left kidney tumor tissue removed from a 43-year-old Caucasian male during nephroureterectomy, regional lymph node excision, and unilateral left adrenalectomy. Pathology indicated a grade 2 renal cell carcinoma in the left kidney. Family history included atherosclerotic coronary artery disease.
LIVRTUE01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from liver tumor tissue removed from a 72-year-old Caucasian male during partial hepatectomy. Pathology indicated metastatic grade 2 (of 4) neuroendocrine carcinoma forming a mass. The patient presented with metastatic liver cancer. Patient history included benign hypertension, type I diabetes, prostatic hyperplasia, prostate cancer, alcohol abuse in remission, and tobacco abuse in remission. Previous surgeries included destruction of a pancreatic lesion, closed prostatic biopsy, transurethral prostatectomy, removal of bilateral testes and total splenectomy. Patient medications included Eulexin, Hytrin, Proscar, Ecotrin, and insulin. Family history included atherosclerotic coronary artery disease and acute myocardial infarction in the mother; atherosclerotic coronary artery disease and type II diabetes in the father.

Table 6

Library	Vector	Library Description
LIVRTUT04	pINCY	Library was constructed using RNA isolated from liver tumor tissue removed from a 50-year-old Caucasian male during a partial hepatectomy. Pathology indicated a grade 3-4 hepatoma, forming a mass. Patient history included benign hypertension and hepatitis. Hepatitis B core antigen and hepatitis B surface antigen was present in the patient.
LUNGFET03	pINCY	Library was constructed using RNA isolated from lung tissue removed from a Caucasian female fetus, who died at 20 weeks' gestation.
LUNGNON07	pINCY	This normalized lung tissue library was constructed from 5.1 million independent clones from a lung tissue library. Starting RNA was made from RNA isolated from lung tissue. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
LUNPTMK03	PSPORT1	This amplified and normalized library was constructed using RNA isolated from pleural tissue removed from a 58-year-old Caucasian female during segmental lung resection. Pathology indicated the pleura consisted of dense connective tissue, with no evidence of tumor. Pathology for the associated tumor tissue indicated metastatic grade 4 leiomyosarcoma, forming a mass in the left lower lobe lung, with extension into the lumen of the pulmonary vein. The patient presented with a malignant retroperitoneum neoplasm with metastasis to lung, an unspecified respiratory abnormality, cough, and died during hospitalization from a tumor embolus. Patient history included hyperlipidemia, paralytic polio, benign bladder neoplasm, normal delivery, benign hypertension, and tobacco abuse in remission. Previous surgeries included adenotonsillectomy, varicose vein ligation and stripping, appendectomy, total abdominal hysterectomy, bilateral salpingo-oophorectomy, and exploratory laparotomy. The patient received radiation therapy for 5.4 weeks. Patient medications included Premarin, Zestril, Butalbital Compound, Centrum vitamins, calcium, amitriptyline, losetolol, and chemotherapy (DTIC, MITO, Adriamycin, cisplatin, GM-CSF, ifosfamide, and VP-16). Family history included benign hypertension and hyperlipidemia in the mother; benign hypertension, hyperlipidemia and skin cancer in the father; hyperlipidemia in the sibling(s); and benign hypertension and cerebrovascular disease in the grandparent(s). Independent clones from this amplified library were normalized in one round using conditions adapted Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791.

Table 6

Library	Vector	Library Description
MCLDTXN03	pINCY	This normalized dendritic cell library was constructed from one million independent clones from a pool of two derived dendritic cell libraries. Starting libraries were constructed using RNA isolated from untreated and treated derived dendritic cells from umbilical cord blood CD34+ precursor cells removed from a male. The cells were derived with granulocyte/macrophage colony stimulating factor (GM-CSF), tumor necrosis factor alpha (TNF alpha), and stem cell factor (SCF). The GM-CSF was added at time 0 at 100 ng/ml, the TNF alpha was added at time 0 at 2.5 ng/ml, and the SCF was added at time 0 at 25 ng/ml. Incubation time was 13 days. The treated cells were then exposed to phorbol myristate acetate (PMA), and Ionomycin. The PMA and Ionomycin were added at 13 days for five hours. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
MENTNOT01	pINCY	Library was constructed using RNA isolated from left tibial meniscus tissue removed from a 16-year-old Caucasian male during a partial left tibial osteotomy with free skin graft. On microscopic exam, this tissue was found to be benign. Pathology for the associated tumor, situated within the proximal 7 cm of the left tibia, indicated metastatic alveolar rhabdomyosarcoma. Patient history included an abnormality of the red blood cells. Family history included osteoarthritis.
MUSCDMT01	pINCY	The library was constructed using RNA isolated from muscle tissue removed from the calf of a 67-year-old Caucasian male during a below the knee amputation and dialysis arteriovenostomy. Pathology indicated multiple necrotic gangrenous areas in all five toes, an area on the medial aspect of the leg at an old incision scar, and an area on the heel of the foot. The vessels showed grade 4 atherosclerosis. The patient presented with hereditary peripheral neuropathy, diabetic neuropathy, deficiency anemia and an unspecified circulatory disease. Patient history included gout, type II diabetes, hyperlipidemia, psoriasis, chronic renal failure, benign hypertension, acute myocardial infarction, and atherosclerotic coronary artery disease. Family history included type II diabetes, acute myocardial infarction, cerebrovascular disease, and nodular lymphoma.
NGANNOT01	PSPORT1	Library was constructed using RNA isolated from tumorous neuroganglion tissue removed from a 9-year-old Caucasian male during a soft tissue excision of the chest wall. Pathology indicated a ganglioneuroma. Family history included asthma.
OVARNOT07	pINCY	Library was constructed using RNA isolated from left ovarian tissue removed from a 28-year-old Caucasian female during a vaginal hysterectomy and removal of the fallopian tubes and ovaries. The tissue was associated with multiple follicular cysts, endometrium in a weakly proliferative phase, and chronic cervicitis of the cervix with squamous metaplasia. Family history included benign hypertension, hyperlipidemia, and atherosclerotic coronary artery disease.

Table 6

Library	Vector	Library Description
PENITUT01	pINCY	Library was constructed using RNA isolated from tumor tissue removed from the penis of a 64-year-old Caucasian male during penile amputation. Pathology indicated a fungating invasive grade 4 squamous cell carcinoma involving the inner wall of the foreskin and extending onto the glans penis. Patient history included benign neoplasm of the large bowel, atherosclerotic coronary artery disease, angina pectoris, gout, and obesity. Family history included malignant pharyngeal neoplasm, chronic lymphocytic leukemia, and chronic liver disease.
PITUNON01	pINCY	This normalized pituitary gland tissue library was constructed from 6.92 million independent clones from a pituitary gland tissue library. Starting RNA was made from pituitary gland tissue removed from a 55-year-old male who died from chronic obstructive pulmonary disease. Neuropathology indicated there were no gross abnormalities, other than mild ventricular enlargement. There was no apparent microscopic abnormality in any of the neocortical areas examined, except for a number of silver positive neurons with apical dendrite staining, particularly in the frontal lobe. The significance of this was undetermined. The only other microscopic abnormality was that there was prominent silver staining with some swollen axons in the CA3 region of the anterior and posterior hippocampus. Microscopic sections of the cerebellum revealed mild Bergmann's gliosis in the Purkinje cell layer. Patient history included schizophrania. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
PLACNOT02	pINCY	Library was constructed using RNA isolated from the placental tissue of a Hispanic female fetus, who was prematurely delivered at 21 weeks' gestation. Serologies of the mother's blood were positive for CMV (cytomegalovirus).
PROSBPS05	pINCY	This subtracted prostate tissue library was constructed using 4.48x10e5 clones from diseased prostate tissue and was subjected to two rounds of subtraction hybridization with 1.56 million clones from a breast tissue library. The starting library for subtraction was constructed using RNA isolated from diseased prostate tissue removed from a 70-year-old Caucasian male during a radical prostatectomy and closed prostatic biopsy. Pathology indicated benign prostatic hypertrophy. Pathology for the matched tumor tissue indicated adenocarcinoma. The patient presented with elevated prostate specific antigen and induration. Patient history included benign hypertension, gastrointestinal bleed, cardiac dysrhythmia, cardiac arrest, hyperlipidemia, alcohol abuse and fractured mandible. Previous surgeries included splenectomy, cholecystectomy and inguinal hernia repair.

Table 6

Library	Vector	Library Description
		Patient medications included Verapamil and antacids. Family history included benign hypertension, myocardial infarction and coronary atherosclerosis in the mother; tobacco abuse and lung cancer in the father; tobacco abuse, cerebrovascular accident and lung cancer in the sibling(s). The hybridization probe for subtraction was derived from a similarly constructed library from RNA isolated from nontumorous breast tissue from a different donor. Subtractive hybridization conditions were based on the methodologies of Swaroop et al., NAR 19 (1991):1954 and Bonaldo, et al. Genome Research 6 (1996): 791.
PROSNOT20	pINCY	Library was constructed using RNA isolated from diseased prostate tissue removed from a 65-year-old Caucasian male during a radical prostatectomy. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated an adenocarcinoma.
PROSTUT05	PSPORT1	Library was constructed using RNA isolated from prostate tumor tissue removed from a 69-year-old Caucasian male during a radical prostatectomy. Pathology indicated adenocarcinoma (Gleason grade 3+4). Adenofibromatous hyperplasia was also present. Family history included congestive heart failure, multiple myeloma, hyperlipidemia, and rheumatoid arthritis.
PROSTUT17	pINCY	The library was constructed using RNA isolated from prostate tumor tissue removed from a 66-year-old Caucasian male during radical prostatectomy, regional lymph node excision, and prostate needle biopsy. Pathology indicated adenocarcinoma Gleason grade 3+4, forming a predominant mass involving the right lobe and the left side centrally. The patient presented with elevated prostate specific antigen (PSA) and induration. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, type II diabetes, hyperlipidemia, and Jakob-Creutzfeldt disease.
SINITMC01	pINCY	This large size-fractionated library was constructed using pooled cDNA from two donors. cDNA was generated using mRNA isolated from ileum tissue removed from a 30-year-old Caucasian female (donor A) during partial colectomy, open liver biopsy, and permanent colostomy, and from ileum tissue removed from a 70-year-old Caucasian female (donor B) during right hemicolectomy, open liver biopsy, sigmoidoscopy, and permanent colostomy. Pathology for the matched tumor tissue (donor A) indicated carcinoid tumor (grade 1 neuroendocrine carcinoma) arising in the terminal ileum. The tumor permeated through the ileal wall into the mesenteric fat and extended into the adherent cecum, where tumor extended through the bowel wall up to the mucosal surface. Multiple lymph nodes were positive for tumor. Additional (2) lymph nodes were also involved by direct tumor extension. Pathology for donor B indicated a non-tumorous margin of ileum. Pathology for the matched tumor (donor B) indicated invasive grade 2 adenocarcinoma forming an ulcerated mass, situated distal to the ileocecal valve.

Table 6

Library	Vector	Library Description
		The tumor invaded through the muscularis propria just into the serosal adipose tissue. One regional lymph node was positive for a microfocus of metastatic adenocarcinoma. Donor A presented with flushing and unspecified abdominal/pelvic symptoms. Patient history included endometriosis, and tobacco and alcohol abuse. Donor B's history included a malignant breast neoplasm, type II diabetes, hyperlipidemia, viral hepatitis, an unspecified thyroid disorder, osteoarthritis, and a malignant skin neoplasm. Donor B's medication included tamoxifen.
SINTNOR01	PCDNA2.1	This random primed library was constructed using RNA isolated from small intestine tissue removed from a 31-year-old Caucasian female during Roux-en-Y gastric bypass. Patient history included clinical obesity.
SMCCNON03	pINCY	This normalized smooth muscle cell library was constructed from 7.56 x 10 <sup>6</sup> independent clones from the SMCCNOT01 library. Starting RNA was made from smooth muscle cell tissue removed from the coronary artery of a 3-year-old Caucasian male. The normalization and hybridization conditions were adapted from Soares et al., (PNAS (1994) 91:9228-9232); Swaroop et al., (NAR (1991) 19:1954); and Bonaldo et al., (Genome Research (1996) 6:791-806), using a significantly longer (48 hour) reannealing hybridization period.
THYMDIT01	pINCY	The library was constructed using RNA isolated from diseased thymus tissue removed from a 16-year-old Caucasian female during a total excision of thymus and regional lymph node excision. Pathology indicated thymic follicular hyperplasia. The right lateral thymus showed reactive lymph nodes. A single reactive lymph node was also identified at the inferior thymus margin. The patient presented with myasthenia gravis, malaise, fatigue, dysphagia, severe muscle weakness, and prominent eyes. Patient history included frozen face muscles. Family history included depressive disorder, hepatitis B, myocardial infarction, atherosclerotic coronary artery disease, leukemia, multiple sclerosis, and lupus.
THYRNOT03	pINCY	Library was constructed using RNA isolated from thyroid tissue removed from the left thyroid of a 28-year-old Caucasian female during a complete thyroidectomy. Pathology indicated a small nodule of adenomatous hyperplasia present in the left thyroid. Pathology for the associated tumor tissue indicated dominant follicular adenoma, forming a well-encapsulated mass in the left thyroid.
TLYMNOT08	pINCY	The library was constructed using RNA isolated from anergocallogenic T-lymphocyte tissue removed from an adult (40-50-year old) Caucasian male. The cells were incubated for 3 days in the presence of 1 microgram/ml OKT3 mAb and 5% human serum.

Table 6

Library	Vector	Library Description
TNFRDNDV01	pCR2-TopoTA	<p>Library was constructed using pooled cDNA from different donors. cDNA was generated using mRNA isolated from pooled small intestine tissue removed from a Caucasian male fetus (donor A) who died at 23 weeks' gestation from premature birth; from lung tissue removed from a Caucasian male fetus (donor B) who died from fetal demise; from pleura tumor tissue removed from a 55-year-old Caucasian female (donor C) during a complete pneumonectomy; from frontal/parietal brain tumor tissue removed from a 2-year-old Caucasian female (donor D) during excision of cerebral meningeal lesion; from liver tumor tissue removed from a 72-year-old Caucasian male (donor E) during partial hepatectomy; from pooled fetal brain tissue removed from a Caucasian male fetus (donor F) who was stillborn with a hypoplastic left heart at 23 weeks' gestation and from brain tissue removed from a Caucasian male fetus (donor G), who died at 23 weeks' gestation from premature birth; from pooled fetal kidney tissue removed from 59, 20-33-week-old male and female fetuses who died from spontaneous abortion; from pooled thymus tissue removed from 9, 18-32-year-old male and female donors who died from sudden death; and from pooled fetal liver tissue removed from 32, 18-24-week-old male and female fetuses. For donor A, serologies were negative. Family history included diabetes in the mother. For donor B, Serologies were negative. For donor C, pathology indicated grade 3 sarcoma most consistent with leiomyosarcoma, uterine primary, forming a bossellated mass replacing the right lower lobe and a portion of the middle lobe. Multiple nodules comprising the tumor show near total necrosis. Smooth muscle actin was positive. Estrogen receptor was negative and progesterone receptor was positive. The patient presented with shortness of breath. Patient history included peptic ulcer disease, normal delivery, anemia, and tobacco abuse in remission. Previous surgeries included total abdominal hysterectomy, bilateral salpingo-oophorectomy, hemorrhoidectomy, endoscopic excision of lung lesion, and appendectomy. Patient medications included Megace, tamoxifen, and Pepcid. Family history included multiple sclerosis in the mother; atherosclerotic coronary artery disease and type II diabetes in the father;</p> <p>and breast cancer in the grandparent(s). For donor D, pathology indicated neuroectodermal tumor with advanced ganglionic differentiation. The lesion was only moderately cellular but was mitotically active with a high MIB-1 labelling index. Neuronal differentiation was widespread and advanced. Multinucleate and dysplastic-appearing forms were readily seen. The glial element was less prominent. The patient presented with motor seizures. Family history included hypertension in the grandparent(s). For donor E, pathology indicated metastatic grade 2 (of 4) neuroendocrine carcinoma forming a mass. The patient presented with metastatic liver cancer. Patient history included benign hypertension, type I diabetes, prostatic hyperplasia, prostate cancer, alcohol abuse in remission, and tobacco abuse in remission. Previous surgeries included destruction of a pancreatic lesion, closed prostatic biopsy, transurethral prostatectomy, removal of bilateral testes and total splenectomy. Patient medications included Eulexin, Hytrin, Proscar, Ecotrin, and insulin. Family history included atherosclerotic coronary artery disease and acute myocardial infarction in the mother;</p>



Table 6

Library	Vector	Library Description
		atherosclerotic coronary artery disease and type II diabetes in the father. For donor F and G, Serologies were negative for both donors and family history for donor G included diabetes in the mother.
TONGTUT01	PSPORT1	Library was constructed using RNA isolated from tongue tumor tissue obtained from a 36-year-old Caucasian male during a hemiglossectomy. Pathology indicated recurrent invasive grade 2 squamous-cell carcinoma.
UTREDMT07	pINCY	Library was constructed using polyA RNA isolated from endometrial tissue removed from a 32-year-old female. The endometrium was in secretory phase and the myometrium was without diagnostic abnormality.
UTRSTMC01	PSPORT1	This large size fractionated library was constructed using pooled cDNA from two donors. cDNA was generated using mRNA isolated from uterus tissue removed from a 49-year-old Caucasian female (donor A) during vaginal hysterectomy and bilateral salpingo-oophorectomy and from uterus tissue removed from a 55-year-old Caucasian female (donor B) during vaginal hysterectomy and bilateral salpingo-oophorectomy. For donor A, pathology indicated inactive endometrium and cervix with no diagnostic changes. Pathology for the matched tumor tissue indicated multiple (6) intramural leiomyomata. The patient presented with excessive menstruation, deficiency anemia, and dysmenorrhea. Patient history included abdominal pregnancy, headache, and chronic obstructive asthma. Previous surgeries included hemorrhoidectomy, knee ligament repair, and intranasal lesion destruction. Patient medications included Azmacort, Proventil, Trazadone, Zostrix HP, iron, Premarin, and vitamin C. Family history included alcohol abuse, atherosclerotic coronary artery disease, upper lobe lung cancer, and carotid endarterectomy in the father; breast fibroadenosis in the sibling(s); and acute myocardial infarction, liver cancer,

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value = 1.0E-8 or less; Full Length sequences: Probability value = 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value = 1.06E-6; Assembled ESTs: fasta Identity = 95% or greater and Match length = 200 bases or greater; fastx E value = 1.0E-8 or less; Full Length sequences: fastx score = 100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value = 1.0E-3 or less

Table 7

Program	Description	Reference	Parameter Threshold
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM, INCY, SMART and TIGRFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	PFAM, INCY, SMART or TIGRFAM hits: Probability value = 1.0E-3 or less; Signal peptide hits: Score = 0 or greater
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score $\geq$ GCG specified "HIGH" value for that particular Prosite motif. Generally, score = 1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score = 120 or greater; Match length = 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score = 3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	

Table 7

Program	Description	Reference	Parameter Threshold
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. On Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence (AAAI) Press, Menlo Park, CA, and MIT Press, Cambridge, MA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

Table 8

SEQ ID NO.	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
102	1333949	6785438H1	SNP00119256	303	312	A	A	G	A51	n/a	n/a	n/a	n/a
103	7035533	1773327H1	SNP00066490	112	2021	T	T	C	I481	n/d	n/d	n/d	n/d
105	2820152	3598946H1	SNP00125766	169	441	A	G	A	K117	n/a	n/a	n/a	n/a
105	2820152	4829541H1	SNP00125767	225	478	C	T	C	S129	n/a	n/a	n/a	n/a
105	2820152	2820152H1	SNP00151315	10	1044	A	A	C	noncoding	n/a	n/a	n/a	n/a
106	2959305	1757802H1	SNP00012320	103	4417	A	A	G	noncoding	n/d	n/a	n/a	n/a
106	2959305	6409880H1	SNP00012320	99	4417	A	A	G	noncoding	n/d	n/a	n/a	n/a
106	2959305	5628628H1	SNP00045303	109	4123	A	A	G	noncoding	n/a	n/a	n/a	n/a
106	2959305	7622357H1	SNP00095526	460	985	G	G	A	T130	0.30	0.15	0.25	0.40
106	2959305	3388052H1	SNP00115018	56	3133	T	G	T	S846	0.28	0.17	0.39	0.26
109	7506225	1978014H1	SNP00036456	103	176	G	G	C	D45	n/d	n/d	n/d	n/d
109	7506225	1212866H1	SNP00041967	111	433	G	G	A	Q130	n/d	n/a	n/a	n/a
109	7506225	1978511H1	SNP00041968	15	491	C	C	A	noncoding	n/a	n/a	n/a	n/a
110	7506227	2240628H1	SNP00042297	106	436	T	T	C	F84	n/a	n/a	n/a	n/a
112	2633315	6578128H1	SNP00026867	103	1546	C	C	T	L480	n/a	n/a	n/a	n/a
112	2633315	6578128H1	SNP00120180	131	1574	G	G	A	S489	n/a	n/a	n/a	n/a
114	045680	6436978H1	SNP00024162	104	2130	C	G	C	F665	n/d	n/a	n/a	n/a
114	045680	7216912H1	SNP00024163	119	2328	T	T	C	T731	0.60	n/a	n/a	n/a
114	045680	7216912H1	SNP00024164	199	2408	T	T	C	V758	n/d	n/d	n/d	n/d
114	045680	2720229H1	SNP00024165	195	2605	A	G	A	R824	n/d	n/d	n/d	n/d
114	045680	2720229H1	SNP00122845	20	2430	G	G	A	E765	n/a	n/a	n/a	n/a
115	1503172	5812127H1	SNP00128531	30	7	C	A	C	noncoding	n/a	n/a	n/a	n/a
116	1818665	7057085H1	SNP00122809	10	3830	G	G	T	Q1204	n/d	n/d	0.98	0.97
116	1818665	1238764H1	SNP00122810	113	3355	T	T	C	L1046	n/d	n/d	n/d	n/d
116	1818665	7586382H2	SNP00122811	195	2165	A	G	A	V649	0.49	n/a	n/a	n/a
116	1818665	6969568H1	SNP00122812	25	1365	G	A	G	G383	n/d	n/a	n/a	n/a
116	1818665	2434340H1	SNP00125177	61	232	G	A	G	R5	n/a	n/a	n/a	n/a
117	3251352	7645422H1	SNP00113283	215	764	G	G	T	L43	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CBI SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
118	55091643	265954H1	SNP00027032	117	205	G	G	C	R46	n/d	n/d	n/d	n/d
120	7506350	5781667H1	SNP00025855	101	156	G	G	A	noncoding	n/a	n/a	n/a	n/a
120	7506350	3746432H1	SNP00025856	10	58	C	C	T	noncoding	0.82	n/d	n/a	n/d
120	7506350	764446H1	SNP00076724	159	3030	T	T	C	noncoding	n/a	n/a	n/a	n/a
120	7506350	7064287H1	SNP00076725	128	1375	T	T	C	L397	n/d	n/d	n/a	0.98
120	7506350	4667308H1	SNP00076726	141	761	G	G	A	Q192	n/a	n/a	n/a	n/a
120	7506350	5524240H1	SNP00117096	124	2913	G	G	A	noncoding	n/a	n/a	n/a	n/a
120	7506350	3594240H1	SNP00117097	142	597	C	C	T	P138	n/d	n/d	n/d	n/d
120	7506350	2172884H1	SNP00138371	221	300	C	C	G	R39	n/a	n/a	n/a	n/a
121	7508370	5812127H1	SNP00128531	30	85	C	A	C	noncoding	n/a	n/a	n/a	n/a
122	2894093	7275356H2	SNP00094396	386	569	G	A	G	S22	n/a	n/a	n/a	n/a
123	7507335	3800593H1	SNP00041736	163	2020	C	T	C	L456	n/a	n/a	n/a	n/a
123	7507335	7218680H1	SNP00132675	107	2695	A	A	G	S681	n/a	n/a	n/a	n/a
123	7507335	1710305H1	SNP00144769	115	2214	T	T	G	G520	n/a	n/a	n/a	n/a
123	7507335	8076822J1	SNP00146274	137	1891	G	G	A	D413	n/a	n/a	n/a	n/a
124	7509081	3404042H1	SNP00023267	145	189	C	C	T	P50	n/d	n/d	n/d	n/d
125	7502450	3168546H1	SNP00038011	108	286	G	G	A	E14	n/a	n/a	n/a	n/a
125	7502450	1456256H1	SNP00045669	102	1124	G	G	C	S293	n/d	n/a	n/a	n/a
125	7502450	3641245H1	SNP00067609	102	720	G	G	T	V158	n/a	n/a	n/a	n/a
125	7502450	6462634H1	SNP00067610	171	956	C	C	T	P237	n/d	n/d	n/d	n/d
126	7501405	7700264J1	SNP00098625	369	333	C	C	A	noncoding	n/a	n/a	n/a	n/a
126	7501405	3150061H1	SNP00102385	173	167	G	T	G	V29	0.91	0.88	0.83	0.93
127	7504528	7610306J1	SNP00013873	26	99	A	G	A	noncoding	0.28	0.62	0.13	0.28
127	7504528	2130414H1	SNP00131705	60	1108	C	C	A	noncoding	n/a	n/a	n/a	n/a
127	7504528	775507H1	SNP00147746	106	729	G	G	A	P147	n/a	n/a	n/a	n/a
129	7509086	4549017H1	SNP00039628	121	514	T	C	T	C156	n/d	n/a	n/a	n/a
129	7509086	509134H1	SNP00131021	163	1122	C	C	T	noncoding	n/a	n/a	n/a	n/a
131	5606114	3438440H1	SNP00000572	41	674	T	T	C	F197	n/a	n/a	n/a	n/a

Table 8

SEQ ID NO:	PID	EST ID	SNP ID	EST SNP	CB1 SNP	EST Allele	Allele 1	Allele 2	Amino Acid	Caucasian Allele 1 frequency	African Allele 1 frequency	Asian Allele 1 frequency	Hispanic Allele 1 frequency
132	7503282	3351651H1	SNP00008796	21	974	T	C	T	noncoding	0.96	n/a	n/a	n/a
132	7503282	5570136H1	SNP00036397	114	1202	G	A	G	noncoding	n/a	n/a	n/a	n/a
133	7503284	1923316H1	SNP00008796	245	1136	C	C	T	noncoding	0.96	n/a	n/a	n/a
134	7510501	3438440H1	SNP00000572	41	493	T	T	C	F125	n/a	n/a	n/a	n/a
135	7500444	2203646H1	SNP00106483	14	609	A	A	C	noncoding	n/d	n/d	n/d	n/d
135	7500444	3990513H1	SNP00113547	129	372	G	G	A	noncoding	n/a	n/a	n/a	n/a
135	7500444	2369510H1	SNP00142909	37	854	A	A	C	noncoding	n/a	n/a	n/a	n/a
136	7510297	4566345H1	SNP00016653	15	341	C	C	T	noncoding	n/a	n/a	n/a	n/a
136	7510297	6064141H1	SNP00016654	5	767	G	G	A	noncoding	n/a	n/a	n/a	n/a
136	7510297	5528027H1	SNP00016655	20	941	T	T	G	noncoding	n/a	n/a	n/a	n/a
136	7510297	1929619H1	SNP00016656	4	1399	C	C	T	noncoding	n/a	n/a	n/a	n/a
136	7510297	1494530H1	SNP00023734	139	451	A	A	C	noncoding	n/a	n/a	n/a	n/a
136	7510297	7130873H1	SNP00065746	106	1443	G	A	G	noncoding	n/a	n/a	n/a	n/a
136	7510297	836712H1	SNP00132864	35	1153	C	C	T	noncoding	n/a	n/a	n/a	n/a
136	7510297	7668977H1	SNP00152047	226	388	C	T	C	noncoding	n/a	n/a	n/a	n/a
137	7640560	924555H1	SNP00032246	71	1067	G	G	C	G280	n/a	n/a	n/a	n/a
137	7640560	7291192H1	SNP00032247	24	1297	C	C	T	S357	n/d	n/d	n/d	n/d
137	7640560	3720649H1	SNP00032248	47	1957	T	T	G	noncoding	n/a	n/a	n/a	n/a
137	7640560	7291192H1	SNP00142824	41	1314	G	G	A	V363	n/a	n/a	n/a	n/a
138	7506087	1436970H1	SNP00008936	95	1300	T	C	T	T403	0.23	0.05	0.23	0.18
138	7506087	7249880H2	SNP00110694	204	1031	T	C	T	W314	n/d	n/d	n/d	n/d

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:

- a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-9 and SEQ ID NO:12-69,
- b) a polypeptide consisting essentially of an amino acid sequence selected from the group consisting of SEQ ID NO:10-11,
- c) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:2-9, SEQ ID NO:12-13, SEQ ID NO:15-16, SEQ ID NO:18-19, SEQ ID NO:23, SEQ ID NO:25-29, SEQ ID NO:33-35, SEQ ID NO:38-39, SEQ ID NO:42-54, SEQ ID NO:57-64, SEQ ID NO:66-67, and SEQ ID NO:69,
- d) a polypeptide comprising a naturally occurring amino acid sequence at least 92% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:20, and SEQ ID NO:55,
- e) a polypeptide comprising a naturally occurring amino acid sequence at least 96% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:14, and SEQ ID NO:37,
- f) a polypeptide comprising a naturally occurring amino acid sequence at least 94% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:17, and SEQ ID NO:56,
- g) a polypeptide comprising a naturally occurring amino acid sequence at least 91% identical to the amino acid sequence of SEQ ID NO:24,
- h) a polypeptide comprising a naturally occurring amino acid sequence at least 93% identical to the amino acid sequence of SEQ ID NO:68,
- i) a polypeptide consisting essentially of a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:21-22, SEQ ID NO:30-32, SEQ ID NO:40-41, and SEQ ID NO:65,
- j) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, and
- k) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-69.

2. An isolated polypeptide of claim 1 selected from the group consisting of:



- a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-9 and SEQ ID NO:12-69, and
- b) a polypeptide consisting essentially of an amino acid sequence selected from the group consisting of SEQ ID NO:10-11.

5

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

10

5. An isolated polynucleotide of claim 4 comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:70-138.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

15

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

20

9. A method of producing a polypeptide of claim 1, the method comprising:

- a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
- b) recovering the polypeptide so expressed.

25

10. A method of claim 9, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-69.

30

11. An isolated antibody which specifically binds to a polypeptide of claim 1.

12. An isolated polynucleotide selected from the group consisting of:

- a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:70-138,

- b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:70-124, SEQ ID NO:126-127, SEQ ID NO:129, SEQ ID NO:131-134, and SEQ ID NO:136-138,
- 5 c) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 97% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:125 and SEQ ID NO:130,
- d) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 91% identical to the polynucleotide sequence of SEQ ID NO:128,
- 10 e) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 99% identical to the polynucleotide sequence of SEQ ID NO:135,
- f) a polynucleotide complementary to a polynucleotide of a),
- g) a polynucleotide complementary to a polynucleotide of b),
- h) a polynucleotide complementary to a polynucleotide of c),
- 15 i) a polynucleotide complementary to a polynucleotide of d),
- j) a polynucleotide complementary to a polynucleotide of e), and
- k) an RNA equivalent of a)-j).

13. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a  
20 polynucleotide of claim 12.

14. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

- 25 a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and
- b) detecting the presence or absence of said hybridization complex, and, optionally, if  
30 present, the amount thereof.

15. A method of claim 14, wherein the probe comprises at least 60 contiguous nucleotides.

16. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
- 5 b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

17. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

10

18. A composition of claim 17, wherein the polypeptide is selected from the group consisting of:

- a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-9 and SEQ ID NO:12-69, and
- 15 b) a polypeptide consisting essentially of an amino acid sequence selected from the group consisting of SEQ ID NO:10-11.

19. A method for treating a disease or condition associated with decreased expression of functional MDDT, comprising administering to a patient in need of such treatment the composition of  
20 claim 17.

20. A method of screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- 25 b) detecting agonist activity in the sample.

21. A composition comprising an agonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

30 22. A method for treating a disease or condition associated with decreased expression of functional MDDT, comprising administering to a patient in need of such treatment a composition of claim 21.

23. A method of screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting antagonist activity in the sample.

5

24. A composition comprising an antagonist compound identified by a method of claim 23 and a pharmaceutically acceptable excipient.

25. A method for treating a disease or condition associated with overexpression of functional MDDT, comprising administering to a patient in need of such treatment a composition of claim 24.

10

26. A method of screening for a compound that specifically binds to the polypeptide of claim 1, the method comprising:

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

15

27. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, the method comprising:

20

- a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

25

30

28. A method of screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- b) detecting altered expression of the target polynucleotide, and
- c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

5

29. A method of assessing toxicity of a test compound, the method comprising:

- a) treating a biological sample containing nucleic acids with the test compound,
- b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,
- c) quantifying the amount of hybridization complex, and
- d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

10

15

20

30. A method for a diagnostic test for a condition or disease associated with the expression of MDDT in a biological sample, the method comprising:

- a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and
- b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

25

31. The antibody of claim 11, wherein the antibody is:

- a) a chimeric antibody,
- b) a single chain antibody,
- c) a Fab fragment,
- d) a F(ab')<sub>2</sub> fragment, or
- e) a humanized antibody.

30

35

32. A composition comprising an antibody of claim 11 and an acceptable excipient.

33. A method of diagnosing a condition or disease associated with the expression of MDDT in a subject, comprising administering to said subject an effective amount of the composition of claim 32.

5           34. A composition of claim 32, further comprising a label.

35. A method of diagnosing a condition or disease associated with the expression of MDDT in a subject, comprising administering to said subject an effective amount of the composition of claim 34.

10

36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:

- 15           a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibodies from the animal, and
- c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-69.

20

37. A polyclonal antibody produced by a method of claim 36.

38. A composition comprising the polyclonal antibody of claim 37 and a suitable carrier.

25           39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:

- 30           a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-69, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibody producing cells from the animal,
- c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,
- d) culturing the hybridoma cells, and

- e) isolating from the culture monoclonal antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-69.

5 40. A monoclonal antibody produced by a method of claim 39.

41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.

10 42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.

43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant immunoglobulin library.

15 44. A method of detecting a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-69 in a sample, the method comprising:

- 20 a) incubating the antibody of claim 11 with the sample under conditions to allow specific binding of the antibody and the polypeptide, and  
b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-69 in the sample.

45. A method of purifying a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-69 from a sample, the method comprising:

- 25 a) incubating the antibody of claim 11 with the sample under conditions to allow specific binding of the antibody and the polypeptide, and  
b) separating the antibody from the sample and obtaining the purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-69.

30 46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.

35 47. A method of generating an expression profile of a sample which contains polynucleotides, the method comprising:

- a) labeling the polynucleotides of the sample,
- b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and
- 5 c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous  
10 nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.  
15

50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is  
20 completely complementary to said target polynucleotide.

52. An array of claim 48, which is a microarray.

53. An array of claim 48, further comprising said target polynucleotide hybridized to a  
25 nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

30 55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.

35



56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

5 58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

10

61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

15

63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

20

65. A polypeptide of claim 1, consisting essentially of the amino acid sequence of SEQ ID NO:10.

66. A polypeptide of claim 1, consisting essentially of the amino acid sequence of SEQ ID NO:11.

25

67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.

68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.

69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.

30

70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.

71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.

35

72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.

73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.

74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.

5 75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.

76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.

77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.

10

78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.

79. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.

15

80. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.

81. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.

82. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.

20

83. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.

84. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.

25

85. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.

86. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:31.

87. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:32.

30

88. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:33.

89. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:34.

35

90. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:35.

91. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:36.

92. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:37.

5 93. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:38.

94. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:39.

95. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:40.

10

96. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:41.

97. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:42.

15

98. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:43.

99. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:44.

100. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:45.

20

101. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:46.

102. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:47.

25

103. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:48.

104. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:49.

105. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:50.

30

106. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:51.

107. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:52.

35

108. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:53.

109. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:54.

110. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:55.

5 111. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:56.

112. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:57.

113. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:58.

10 114. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:59.

115. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:60.

15 116. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:61.

117. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:62.

118. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:63.

20 119. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:64.

120. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:65.

25 121. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:66.

122. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:67.

123. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:68.

30 124. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:69.

125. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:70.

35

126. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:71.

127. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:72.

128. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:73.

129. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:74.

130. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:75.

131. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:76.

132. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:77.

133. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:78.

134. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:79.

135. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:80.

136. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:81.

137. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:82.

138. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:83.

139. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
5 NO:84.

140. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:85.

141. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
10 NO:86.

142. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:87.

15 143. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:88.

144. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
20 NO:89.

145. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:90.

25 146. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:91.

147. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:92.

30 148. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:93.

149. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
35 NO:94.

150. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:95.

151. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
5 NO:96.

152. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:97.

10 153. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:98.

154. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:99.

15 155. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:100.

156. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
20 NO:101.

157. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:102.

25 158. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:103.

159. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:104.

30 160. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:105.

161. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
35 NO:106.

162. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:107.

5 163. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:108.

164. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:109.

10 165. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:110.

166. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:111.

15 167. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:112.

20 168. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:113.

169. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:114.

25 170. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:115.

171. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:116.

30 172. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:117.

35 173. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:118.



174. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:119.

175. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:120.

176. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:121.

177. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:122.

178. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:123.

179. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:124.

180. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:125.

181. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:126.

182. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:127.

183. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:128.

184. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:129.

185. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:130.

186. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:131.

187. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
5 NO:132.

188. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:133.

10 189. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:134.

190. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:135.

15 191. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:136.

192. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
20 NO:137.

193. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:138.

25

<110> INCYTE GENOMICS, INC.  
TANG, Y. Tom  
CHAWLA, Narinder K.  
LU, Dyung Aina M.  
KHAN, Farrah A.  
GANDHI, Ameena R.  
SWARNAKAR, Anita  
AZIMZAI, Yalda  
MARQUIS, Joseph P.  
SPRAGUE, William W.  
EMERLING, Brooke M.  
YUE, Henry  
BOROWSKY, Mark L.  
BECHA, Shanya D.  
ISON, Craig H.  
ELLIOTT, Vicki S.  
HAFALIA, April J.A.  
RING, Huijun Z.  
WARREN, Bridget A.  
GIETZEN, Kimberly J.  
TRAN, Uyen K.  
LEE, Soo Yeun  
LEE, Ernestine A.  
RICHARDSON, Thomas W.  
KABLE, Amy E.  
BURFORD, Neil  
LEHR-MASON, Patricia M.  
GORVAD, Ann E.  
LEE, Sally  
BLAKE, Julie J.  
HONCHELL, Cynthia D.  
THANGAVELU, Kavitha  
RAMKUMAR, Jayalaxmi  
CHIEN, David  
JIN, Pei  
CHANG, Hsin-Ru  
BAUGHN, Mariah R.  
NGUYEN, Danniel B.  
KHARE, Reena  
BHATIA, Umesh  
BURRILL, John D.  
HO, Anne  
ZHENG, Wenjin

<120> MOLECULES FOR DISEASE DETECTION AND TREATMENT

<130> PF-1302 PCT

<140> To Be Assigned

<141> Herewith

<150> US 60/334,182  
<151> 2001-11-28

<150> US 60/342,052  
<151> 2001-12-18

<150> US 60/350,410  
<151> 2002-01-18

<150> US 60/353,284  
<151> 2002-02-01

<150> US 60/363,649  
 <151> 2002-03-08

<160> 138  
 <170> PERL Program

<210> 1  
 <211> 126  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1393336CD1

<400> 1  
 Met Ile Ile Asn Leu Ile Leu Phe Tyr Gly Phe Arg Ser Arg Gln  
 1 5 10 15  
 Pro Ile Ser Ala Pro Leu Phe Ser Cys Pro Asp Lys Asn Lys Val  
 20 25 30  
 Asn Phe Ile Pro Thr Gly Ser Ala Phe Cys Pro Val Lys Leu Leu  
 35 40 45  
 Gly Pro Leu Leu Pro Ala Ser Asp Leu Met Leu Lys Asn Ser Pro  
 50 55 60  
 Asn Ser Gly Gln Ser Ser Ala Leu Ala Thr Leu Thr Val Glu Gln  
 65 70 75  
 Leu Ser Ser Arg Val Ser Phe Thr Ser Leu Ser Asp Asp Thr Ser  
 80 85 90  
 Thr Ala Gly Ser Met Glu Ala Ser Val Gln Gln Pro Ser Gln Gln  
 95 100 105  
 Gln Gln Leu Leu Gln Glu Leu Gln Gly Glu Asp His Ile Ser Ala  
 110 115 120  
 Gln Asn Tyr Val Ile Ile  
 125

<210> 2  
 <211> 109  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1431502CD1

<400> 2  
 Met Ala Pro Leu His Ser Ser Leu Gly Asp Arg Ala Arg Leu Ser  
 1 5 10 15  
 Lys Lys Lys Glu Lys Lys Arg Lys Glu Lys Lys Arg Asn Ile  
 20 25 30  
 Leu Thr Trp Pro Gly Thr Val Ala His Ala Cys Asn Pro Arg Thr  
 35 40 45  
 Leu Gly Cys Gln Gly Arg Arg Ile Ser Lys Leu Gly Val Gln Asp  
 50 55 60  
 Gln Pro Gly Gln His Gly Glu Thr Trp Ser Leu Leu Asn Ile His  
 65 70 75  
 Lys Ile Ser Trp Ala Trp Trp His Pro Pro Val Ile Pro Ala Thr  
 80 85 90  
 Gln Glu Ala Glu Ala Glu Glu Thr Leu Glu Pro Gly Gly Asp Cys  
 95 100 105  
 Pro Thr Ala Phe

<210> 3

<211> 464  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2445220CD1

<400> 3  
 Met Gly Val Ser Lys Leu Asp Ile Leu Tyr Arg Arg Leu Leu Leu  
 1 5 10 15  
 Thr Lys Leu Phe Ile Arg Gly Trp Gly Arg Pro Glu Asp Leu Lys  
 20 25 30  
 Arg Leu Phe Glu Phe Arg Lys Met Ile Gly Asn Arg Glu Arg Cys  
 35 40 45  
 Gln Asn Leu Val Ser Ser Asp Tyr Pro Val His Ile Asp Lys Ile  
 50 55 60  
 Glu Glu Gln Ser Asp Cys Lys Ile Leu Asp Gly His Phe Val Ser  
 65 70 75  
 Pro Met Ala His Tyr Val Pro Asp Ile Met Pro Ile Glu Ser Val  
 80 85 90  
 Ile Ala Arg Phe Gln Phe Ile Val Pro Lys Glu Trp Asn Ser Lys  
 95 100 105  
 Tyr Arg Pro Val Cys Ile His Leu Ala Gly Thr Gly Asp His His  
 110 115 120  
 Tyr Trp Arg Arg Arg Thr Leu Met Ala Arg Pro Met Ile Lys Glu  
 125 130 135  
 Ala Arg Met Ala Ser Leu Leu Leu Glu Asn Pro Tyr Tyr Gly Cys  
 140 145 150  
 Arg Lys Pro Lys Asp Gln Val Arg Ser Ser Leu Lys Asn Val Ser  
 155 160 165  
 Asp Leu Phe Val Met Gly Gly Ala Leu Val Leu Glu Ser Ala Ala  
 170 175 180  
 Leu Leu His Trp Leu Glu Arg Glu Gly Tyr Gly Pro Leu Gly Met  
 185 190 195  
 Thr Gly Ile Ser Met Gly Gly His Met Ala Ser Leu Ala Val Ser  
 200 205 210  
 Asn Trp Pro Lys Pro Met Pro Leu Ile Pro Cys Leu Ser Trp Ser  
 215 220 225  
 Thr Ala Ser Gly Val Phe Thr Thr Gly Val Leu Ser Lys Ser Ile  
 230 235 240  
 Asn Trp Arg Glu Leu Glu Lys Gln Tyr Tyr Thr Gln Thr Val Tyr  
 245 250 255  
 Glu Glu Glu Ile Ile His Met Leu Glu Tyr Cys Gly Thr Asp Ser  
 260 265 270  
 Phe Lys Met Gly Gln Glu Phe Val Lys His Phe Thr Ser Ser Ala  
 275 280 285  
 Asp Lys Leu Thr Asn Leu Asn Leu Val Ser Arg Thr Leu Asn Leu  
 290 295 300  
 Asp Ile Ser Asn Gln Val Val Ser Gln Lys Pro Ala Asp Cys His  
 305 310 315  
 Asn Ser Ser Lys Thr Ser Val Ser Ala Thr Ser Glu Gly Leu Leu  
 320 325 330  
 Leu Gln Asp Thr Ser Lys Met Lys Arg Phe Asn Gln Thr Leu Ser  
 335 340 345  
 Thr Asn Lys Ser Gly Tyr Thr Ser Arg Asn Pro Gln Ser Tyr His  
 350 355 360  
 Leu Leu Ser Lys Glu Gln Ser Arg Asn Ser Leu Arg Lys Glu Ser  
 365 370 375  
 Leu Ile Phe Met Lys Gly Val Met Asp Glu Cys Thr His Val Ala  
 380 385 390  
 Asn Phe Ser Val Pro Val Asp Pro Ser Leu Ile Ile Val Val Gln  
 395 400 405

Ala	Lys	Glu	Asp	Ala	Tyr	Ile	Pro	Arg	Thr	Gly	Val	Arg	Ser	Leu
				410					415					420
Gln	Glu	Ile	Trp	Pro	Gly	Cys	Glu	Ile	Arg	Tyr	Leu	Glu	Gly	Gly
				425					430					435
His	Ile	Ser	Ala	Tyr	Leu	Phe	Lys	Gln	Gly	Leu	Phe	Arg	Gln	Ala
				440					445					450
Ile	Tyr	Asp	Ala	Phe	Asp	Arg	Phe	Leu	His	Lys	Tyr	Ala	Asn	
				455					460					

&lt;210&gt; 4

&lt;211&gt; 273

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 5504385CD1

&lt;400&gt; 4

Met	Glu	Thr	Ser	Ser	Lys	Gly	Leu	Leu	Thr	Gln	Val	Thr	Gln	Phe
1				5					10					15
Trp	Asn	Leu	Leu	Asp	Asp	Leu	Ala	Gln	Ser	Asp	Pro	Glu	Gly	Tyr
				20					25					30
Glu	Lys	Phe	Ile	Gln	Gln	Gln	Leu	Lys	Glu	Gly	Lys	Gln	Leu	Cys
				35					40					45
Ala	Ala	Pro	Glu	Pro	Gln	Leu	Cys	Leu	Gln	Thr	Arg	Ile	Leu	Lys
				50					55					60
Pro	Lys	Glu	Lys	Ile	Leu	Phe	Ile	Asn	Leu	Cys	Gln	Trp	Thr	Arg
				65					70					75
Ile	Pro	Ala	Pro	Gln	Ser	Thr	Thr	His	Pro	Val	Pro	Leu	Thr	Val
				80					85					90
Gly	Lys	Pro	Glu	Asp	Thr	Thr	Glu	Ile	Ser	Asp	Ala	Tyr	Thr	Val
				95					100					105
Ile	Asp	Val	Ala	Tyr	Asn	Pro	Asp	Val	Leu	His	Ala	Ala	Glu	Lys
				110					115					120
Asp	Gln	Val	Lys	Lys	Asn	Gln	Leu	Ile	Gln	Met	Ala	Met	Lys	Cys
				125					130					135
Ile	Glu	Glu	Lys	Phe	Gln	Phe	Thr	Leu	Ser	His	Ser	Tyr	His	Ile
				140					145					150
Thr	Lys	Phe	Arg	Ile	Lys	Gly	Ser	Ile	Gln	Arg	Met	Lys	Gln	Asn
				155					160					165
Leu	Met	Gly	Ile	Gln	Thr	Asp	Ser	Ile	Asp	Leu	Arg	Glu	Lys	Met
				170					175					180
Arg	Arg	Glu	Leu	Thr	Leu	Gly	Gln	Ile	Arg	Ser	Ser	Thr	Met	Ser
				185					190					195
Asn	Pro	Asp	His	Phe	Pro	Gln	Leu	Leu	Leu	Pro	Lys	Asp	Gln	Val
				200					205					210
Ser	Gly	Lys	Ala	Val	Cys	Leu	Ile	Glu	Glu	Ile	Ser	Ser	Thr	Glu
				215					220					225
Ile	Gln	Val	Glu	Met	Lys	Met	Pro	Ala	Tyr	Glu	Leu	Lys	Ile	Val
				230					235					240
His	Asp	His	Ser	Glu	Lys	Pro	Leu	Lys	Ile	Glu	Leu	Lys	Val	Glu
				245					250					255
Leu	Pro	Gly	Ile	Asn	Ser	Val	Ser	Leu	Cys	Asp	Leu	Ser	Val	Ser
				260					265					270

Glu Val Ser

&lt;210&gt; 5

&lt;211&gt; 877

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 6974948CD1

&lt;400&gt; 5

Met	Pro	Ala	Ala	Asn	Met	Met	Glu	Asn	Arg	Pro	Leu	Pro	Ala	Leu	1	5	10	15
Gln	Val	Pro	Glu	Pro	Gln	Gly	Ala	Pro	Glu	Gly	Ser	Pro	Val	Trp	20	25	30	35
Ser	Ser	Ser	Ser	Thr	Pro	Thr	Leu	Arg	Arg	Arg	Arg	Phe	Lys	Met	40	45	50	55
Arg	Arg	Met	Lys	Asn	Val	Gln	Glu	Gln	Ser	Leu	Glu	Ala	Gly	Leu	60	65	70	75
Ala	Arg	Asp	Leu	Pro	Ala	Val	Leu	Ala	Pro	Gly	Lys	Glu	Phe	Leu	80	85	90	95
Gln	Leu	Pro	Ser	Ile	Glu	Ile	Thr	Pro	Ser	Ser	Asp	Glu	Asp	Thr	100	105	110	115
Pro	Trp	Ser	Asn	Cys	Ser	Thr	Pro	Ser	Ala	Ser	Pro	Arg	Arg	Lys	120	125	130	135
Arg	Phe	Leu	Leu	Arg	Lys	Trp	Leu	Arg	Val	Arg	Glu	Arg	Lys	Glu	140	145	150	155
Cys	Ser	Glu	Ser	Ser	Ser	Gln	Gln	Ser	Ser	Gln	Gln	Ser	Ser	His	160	165	170	175
Asp	Asp	Asp	Ser	Ser	Arg	Phe	Leu	Ser	Pro	Arg	Ala	Arg	Glu	Glu	180	185	190	195
Ser	Thr	Ala	Ser	Asn	Ser	Asn	Arg	Ser	Thr	Pro	Ala	Cys	Ser	Pro	200	205	210	215
Ile	Leu	Arg	Lys	Arg	Ser	Arg	Ser	Pro	Thr	Pro	Gln	Asn	Gln	Asp	220	225	230	235
Gly	Asp	Thr	Met	Val	Glu	Lys	Gly	Ser	Asp	His	Ser	Ser	Asp	Lys	240	245	250	255
Ser	Pro	Ser	Thr	Pro	Glu	Gln	Gly	Val	Gln	Arg	Ser	Cys	Ser	Ser	260	265	270	275
Gln	Ser	Gly	Arg	Ser	Gly	Gly	Lys	Asn	Ser	Lys	Lys	Ser	Gln	Ser	280	285	290	295
Trp	Tyr	Asn	Val	Leu	Ser	Pro	Thr	Tyr	Lys	Gln	Arg	Asn	Glu	Asp	300	305	310	315
Phe	Arg	Lys	Leu	Phe	Lys	Gln	Leu	Pro	Asp	Thr	Glu	Arg	Leu	Ile	320	325	330	335
Val	Asp	Tyr	Ser	Cys	Ala	Leu	Gln	Arg	Asp	Ile	Leu	Leu	Gln	Gly	340	345	350	355
Arg	Leu	Tyr	Leu	Ser	Glu	Asn	Trp	Ile	Cys	Phe	Tyr	Ser	Asn	Ile	360	365	370	375
Phe	Arg	Trp	Glu	Thr	Leu	Leu	Thr	Val	Arg	Leu	Lys	Asp	Ile	Cys	380	385	390	395
Ser	Met	Thr	Lys	Glu	Lys	Thr	Ala	Arg	Leu	Ile	Pro	Asn	Ala	Ile	400	405	410	415
Gln	Val	Cys	Thr	Asp	Ser	Glu	Lys	His	Phe	Phe	Thr	Ser	Phe	Gly	420	425	430	435
Ala	Arg	Asp	Arg	Thr	Tyr	Met	Met	Met	Phe	Arg	Leu	Trp	Gln	Asn				
Ala	Leu	Leu	Glu	Lys	Pro	Leu	Cys	Pro	Lys	Glu	Leu	Trp	His	Phe				
Val	His	Gln	Cys	Tyr	Gly	Asn	Glu	Leu	Gly	Leu	Thr	Ser	Asp	Asp				
Glu	Asp	Tyr	Val	Pro	Pro	Asp	Asp	Asp	Phe	Asn	Thr	Met	Gly	Tyr				
Cys	Glu	Glu	Ile	Pro	Val	Glu	Glu	Asn	Glu	Val	Asn	Asp	Ser	Ser				
Ser	Lys	Ser	Ser	Ile	Glu	Thr	Lys	Pro	Asp	Ala	Ser	Pro	Gln	Leu				
Pro	Lys	Lys	Ser	Ile	Thr	Asn	Ser	Thr	Leu	Thr	Ser	Thr	Gly	Ser				

Ser	Glu	Ala	Pro	Val	Ser	Phe	Asp	Gly	Leu	Pro	Leu	Glu	Glu	Glu
				440					445					450
Ala	Leu	Glu	Gly	Asp	Gly	Ser	Leu	Glu	Lys	Glu	Leu	Ala	Ile	Asp
				455					460					465
Asn	Ile	Met	Gly	Glu	Lys	Ile	Glu	Met	Ile	Ala	Pro	Val	Asn	Ser
				470					475					480
Pro	Ser	Leu	Asp	Phe	Asn	Asp	Asn	Glu	Asp	Ile	Pro	Thr	Glu	Leu
				485					490					495
Ser	Asp	Ser	Ser	Asp	Thr	His	Asp	Glu	Gly	Glu	Val	Gln	Ala	Phe
				500					505					510
Tyr	Glu	Asp	Leu	Ser	Gly	Arg	Gln	Tyr	Val	Asn	Glu	Val	Phe	Asn
				515					520					525
Phe	Ser	Val	Asp	Lys	Leu	Tyr	Asp	Leu	Leu	Phe	Thr	Asn	Ser	Pro
				530					535					540
Phe	Gln	Arg	Asp	Phe	Met	Glu	Gln	Arg	Arg	Phe	Ser	Asp	Ile	Ile
				545					550					555
Phe	His	Pro	Trp	Lys	Lys	Glu	Glu	Asn	Gly	Asn	Gln	Ser	Arg	Val
				560					565					570
Ile	Leu	Tyr	Thr	Ile	Thr	Leu	Thr	Asn	Pro	Leu	Ala	Pro	Lys	Thr
				575					580					585
Ala	Thr	Val	Arg	Glu	Thr	Gln	Thr	Met	Tyr	Lys	Ala	Ser	Gln	Glu
				590					595					600
Ser	Glu	Cys	Tyr	Val	Ile	Asp	Ala	Glu	Val	Leu	Thr	His	Asp	Val
				605					610					615
Pro	Tyr	His	Asp	Tyr	Phe	Tyr	Thr	Ile	Asn	Arg	Tyr	Thr	Leu	Thr
				620					625					630
Arg	Val	Ala	Arg	Asn	Lys	Ser	Arg	Leu	Arg	Val	Ser	Thr	Glu	Leu
				635					640					645
Arg	Tyr	Arg	Lys	Gln	Pro	Trp	Gly	Leu	Val	Lys	Thr	Phe	Ile	Glu
				650					655					660
Lys	Asn	Phe	Trp	Ser	Gly	Leu	Glu	Asp	Tyr	Phe	Arg	His	Leu	Glu
				665					670					675
Ser	Glu	Leu	Ala	Lys	Thr	Glu	Ser	Thr	Tyr	Leu	Ala	Glu	Met	His
				680					685					690
Arg	Gln	Ser	Pro	Lys	Glu	Lys	Ala	Ser	Lys	Thr	Thr	Thr	Val	Arg
				695					700					705
Arg	Arg	Lys	Arg	Pro	His	Ala	His	Leu	Arg	Val	Pro	His	Leu	Glu
				710					715					720
Glu	Val	Met	Ser	Pro	Val	Thr	Thr	Pro	Thr	Asp	Glu	Asp	Val	Gly
				725					730					735
His	Arg	Ile	Lys	His	Val	Ala	Gly	Ser	Thr	Gln	Thr	Arg	His	Ile
				740					745					750
Pro	Glu	Asp	Thr	Pro	Asn	Gly	Phe	His	Leu	Gln	Ser	Val	Ser	Lys
				755					760					765
Leu	Leu	Leu	Val	Ile	Ser	Cys	Val	Leu	Val	Leu	Leu	Val	Ile	Leu
				770					775					780
Asn	Met	Met	Leu	Phe	Tyr	Lys	Leu	Trp	Met	Leu	Glu	Tyr	Thr	Thr
				785					790					795
Gln	Thr	Leu	Thr	Ala	Trp	Gln	Gly	Leu	Arg	Leu	Gln	Glu	Arg	Leu
				800					805					810
Pro	Gln	Ser	Gln	Thr	Glu	Trp	Ala	Gln	Leu	Leu	Glu	Ser	Gln	Gln
				815					820					825
Lys	Tyr	His	Asp	Thr	Glu	Leu	Gln	Lys	Trp	Arg	Glu	Ile	Ile	Lys
				830					835					840
Ser	Ser	Val	Met	Leu	Leu	Asp	Gln	Met	Lys	Asp	Ser	Leu	Ile	Asn
				845					850					855
Leu	Gln	Asn	Gly	Ile	Arg	Ser	Arg	Asp	Tyr	Thr	Ser	Glu	Ser	Glu
				860					865					870
Glu	Lys	Arg	Asn	Arg	Tyr	His								
				875										

<210> 6  
<211> 756



<212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7501636CD1

<400> 6  
 Met Gly Ser Thr Asp Ser Lys Leu Asn Phe Arg Lys Ala Val Ile  
   1                  5                  10                  15  
 Gln Leu Thr Thr Lys Thr Gln Pro Val Glu Ala Thr Asp Asp Ala  
                   20                  25                  30  
 Phe Trp Asp Gln Phe Trp Ala Asp Thr Ala Thr Ser Val Gln Asp  
                   35                  40                  45  
 Val Phe Ala Leu Val Pro Ala Ala Glu Ile Arg Ala Val Arg Glu  
                   50                  55                  60  
 Glu Ser Pro Ser Asn Leu Ala Thr Leu Cys Tyr Lys Ala Val Glu  
                   65                  70                  75  
 Lys Leu Val Gln Gly Ala Glu Ser Gly Cys His Ser Glu Lys Glu  
                   80                  85                  90  
 Lys Gln Ile Val Leu Asn Cys Ser Arg Leu Leu Thr Arg Val Leu  
                   95                  100                 105  
 Pro Tyr Ile Phe Glu Asp Pro Asp Trp Arg Gly Phe Phe Trp Ser  
                  110                 115                 120  
 Thr Val Pro Gly Ala Gly Arg Gly Gly Gly Glu Glu Asp Asp  
                  125                 130                 135  
 Glu His Ala Arg Pro Leu Ala Glu Ser Leu Leu Leu Ala Ile Ala  
                  140                 145                 150  
 Asp Leu Leu Phe Cys Pro Asp Phe Thr Val Gln Ser His Arg Arg  
                  155                 160                 165  
 Ser Thr Val Asp Ser Ala Glu Asp Val His Ser Leu Asp Ser Cys  
                  170                 175                 180  
 Glu Tyr Ile Trp Glu Ala Gly Val Gly Phe Ala His Ser Pro Gln  
                  185                 190                 195  
 Pro Asn Tyr Ile His Asp Met Asn Arg Met Glu Leu Leu Lys Leu  
                  200                 205                 210  
 Leu Leu Thr Cys Phe Ser Glu Ala Met Tyr Leu Pro Pro Ala Pro  
                  215                 220                 225  
 Glu Ser Gly Ser Thr Asn Pro Trp Val Gln Phe Phe Cys Ser Thr  
                  230                 235                 240  
 Glu Asn Arg His Ala Leu Pro Leu Phe Thr Ser Leu Leu Asn Thr  
                  245                 250                 255  
 Val Cys Ala Tyr Asp Pro Val Gly Tyr Gly Ile Pro Tyr Asn His  
                  260                 265                 270  
 Leu Leu Phe Ser Asp Tyr Arg Glu Pro Leu Val Glu Glu Ala Ala  
                  275                 280                 285  
 Gln Val Leu Ile Val Thr Leu Asp His Asp Ser Ala Ser Ser Ala  
                  290                 295                 300  
 Ser Pro Thr Val Asp Gly Thr Thr Thr Gly Thr Ala Met Asp Asp  
                  305                 310                 315  
 Ala Asp Pro Pro Gly Pro Glu Asn Leu Phe Val Asn Tyr Leu Ser  
                  320                 325                 330  
 Arg Ile His Arg Glu Glu Asp Phe Gln Phe Ile Leu Lys Gly Ile  
                  335                 340                 345  
 Ala Arg Leu Leu Ser Asn Pro Leu Leu Gln Thr Tyr Leu Pro Asn  
                  350                 355                 360  
 Ser Thr Lys Lys Ile Gln Phe His Gln Glu Leu Leu Val Leu Phe  
                  365                 370                 375  
 Trp Lys Leu Cys Asp Phe Asn Lys Lys Phe Leu Phe Phe Val Leu  
                  380                 385                 390  
 Lys Ser Ser Asp Val Leu Asp Ile Leu Val Pro Ile Leu Phe Phe  
                  395                 400                 405  
 Leu Asn Asp Ala Arg Ala Asp Gln Ser Arg Val Gly Leu Met His

	410		415		420
Ile Gly Val Phe	Ile Leu Leu Leu Leu	Ser Gly Glu Arg Asn	Phe		
	425		430		435
Gly Val Arg Leu	Asn Lys Pro Tyr Ser	Ile Arg Val Pro Met	Asp		
	440		445		450
Ile Pro Val Phe	Thr Gly Thr His Ala	Asp Leu Leu Ile Val	Val		
	455		460		465
Phe His Lys Ile	Ile Thr Ser Gly His	Gln Arg Leu Gln Pro	Leu		
	470		475		480
Phe Asp Cys Leu	Leu Thr Ile Val Val	Asn Val Ser Pro Tyr	Leu		
	485		490		495
Lys Ser Leu Ser	Met Val Thr Ala Asn	Lys Leu Leu His Leu	Leu		
	500		505		510
Glu Ala Phe Ser	Thr Thr Trp Phe Leu	Phe Ser Ala Ala Gln	Asn		
	515		520		525
His His Leu Val	Phe Phe Leu Leu Glu	Val Phe Asn Asn Ile	Ile		
	530		535		540
Gln Tyr Gln Phe	Asp Gly Asn Ser Asn	Leu Val Tyr Ala Ile	Ile		
	545		550		555
Arg Lys Arg Ser	Ile Phe His Gln Leu	Ala Asn Leu Pro Thr	Asp		
	560		565		570
Pro Pro Thr Ile	His Lys Ala Leu Gln	Arg Arg Arg Arg Thr	Pro		
	575		580		585
Glu Pro Leu Ser	Arg Thr Gly Ser Gln	Glu Gly Thr Ser Met	Glu		
	590		595		600
Gly Ser Arg Pro	Ala Ala Pro Ala Glu	Pro Gly Thr Leu Lys	Thr		
	605		610		615
Ser Leu Val Ala	Thr Pro Gly Ile Asp	Lys Leu Thr Glu Lys	Ser		
	620		625		630
Gln Val Ser Glu	Asp Gly Thr Leu Arg	Ser Leu Glu Pro Glu	Pro		
	635		640		645
Gln Gln Ser Leu	Glu Asp Gly Ser Pro	Ala Lys Gly Glu Pro	Ser		
	650		655		660
Gln Ala Trp Arg	Glu Gln Arg Arg Pro	Ser Thr Ser Ser Ala	Ser		
	665		670		675
Gly Gln Trp Ser	Pro Thr Pro Glu Trp	Val Leu Ser Trp Lys	Ser		
	680		685		690
Lys Leu Pro Leu	Gln Thr Ile Met Arg	Leu Leu Gln Val Leu	Val		
	695		700		705
Pro Gln Val Glu	Lys Ile Cys Ile Asp	Lys Gly Leu Thr Asp	Glu		
	710		715		720
Ser Glu Ile Leu	Arg Phe Leu Gln His	Gly Thr Leu Val Gly	Leu		
	725		730		735
Leu Pro Val Pro	His Pro Ile Leu Ile	Arg Met Trp Thr Pro	Leu		
	740		745		750
Ser Gly Thr Thr	Pro Thr				
	755				

<210> 7  
 <211> 363  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2535717CD1

<400> 7  
 Met Asp Ser Phe Gly Gln Pro Arg Pro Glu Asp Asn Gln Ser Val  
 1 5 10 15  
 Val Arg Arg Met Gln Lys Lys Tyr Trp Lys Thr Lys Gln Val Phe  
 20 25 30  
 Ile Lys Ala Thr Gly Lys Lys Glu Asp Glu His Leu Val Ala Ser

	35		40		45									
Asp	Ala	Glu	Leu	Asp	Ala	Lys	Leu	Glu	Val	Phe	His	Ser	Val	Gln
	50				55									60
Glu	Thr	Cys	Thr	Glu	Leu	Leu	Lys	Ile	Ile	Glu	Lys	Tyr	Gln	Leu
	65				70									75
Arg	Leu	Asn	Val	Ile	Ser	Glu	Glu	Glu	Asn	Glu	Leu	Gly	Leu	Phe
	80				85									90
Leu	Lys	Phe	Gln	Ala	Glu	Arg	Asp	Ala	Thr	Gln	Ala	Gly	Lys	Met
	95				100									105
Met	Asp	Ala	Thr	Gly	Lys	Ala	Leu	Cys	Ser	Ser	Ala	Lys	Gln	Arg
	110				115									120
Leu	Ala	Leu	Cys	Thr	Pro	Leu	Ser	Arg	Leu	Lys	Gln	Glu	Val	Ala
	125				130									135
Thr	Phe	Ser	Gln	Arg	Ala	Val	Ser	Asp	Thr	Leu	Met	Thr	Ile	Asn
	140				145									150
Arg	Met	Glu	Gln	Ala	Arg	Thr	Glu	Tyr	Arg	Gly	Ala	Leu	Leu	Trp
	155				160									165
Met	Lys	Asp	Val	Ser	Gln	Glu	Leu	Asp	Pro	Asp	Thr	Leu	Lys	Gln
	170				175									180
Met	Glu	Lys	Phe	Arg	Lys	Val	Gln	Met	Gln	Val	Arg	Asn	Ser	Lys
	185				190									195
Ala	Ser	Phe	Asp	Lys	Leu	Lys	Met	Asp	Val	Cys	Gln	Lys	Val	Asp
	200				205									210
Leu	Leu	Gly	Ala	Ser	Arg	Cys	Asn	Met	Leu	Ser	His	Ser	Leu	Thr
	215				220									225
Thr	Tyr	Gln	Lys	Thr	Ala	Arg	Met	Met	Ser	Gln	Ile	His	Glu	Ala
	230				235									240
Cys	Ile	Gly	Phe	His	Pro	Tyr	Asp	Phe	Val	Ala	Leu	Lys	Gln	Leu
	245				250									255
Gln	Asp	Thr	Pro	Ser	Lys	Ile	Ser	Glu	Asp	Asn	Lys	Asp	Glu	Gln
	260				265									270
Ile	Gly	Gly	Phe	Leu	Thr	Glu	Gln	Leu	Asn	Lys	Leu	Val	Leu	Ser
	275				280									285
Asp	Glu	Glu	Ala	Ser	Phe	Glu	Ser	Glu	Gln	Asp	Trp	Val	Ser	Gln
	290				295									300
Glu	Glu	Ser	Glu	Leu	Cys	Leu	Ser	His	Thr	Asp	Asn	Gln	Pro	Val
	305				310									315
Pro	Ser	Gln	Ser	Pro	Lys	Lys	Leu	Thr	Arg	Ser	Pro	Asn	Asn	Gly
	320				325									330
Asn	Gln	Asp	Met	Ser	Ala	Trp	Phe	Asn	Leu	Phe	Ala	Asp	Leu	Asp
	335				340									345
Pro	Leu	Ser	Asn	Pro	Asp	Ala	Ile	Gly	His	Ser	Asp	Asp	Glu	Leu
	350				355									360

Leu Asn Ala

<210> 8  
 <211> 608  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 6119548CD1

<400> 8  
 Met Leu Ser Leu Cys Ala Gln Asp Leu Thr Gln Met Leu Ala Leu  
 1 5 10 15  
 Ser Arg His Ser Leu Leu Ser Pro Leu Leu Ser Val Thr Ser Phe  
 20 25 30  
 Arg Arg Phe Tyr Arg Gly Asp Ser Pro Thr Asp Ser Gln Lys Asp  
 35 40 45  
 Met Ile Glu Ile Pro Leu Pro Pro Trp Gln Glu Arg Thr Asp Glu

	50		55		60
Ser Ile Glu Thr	Lys Arg Ala Arg Leu	Leu Tyr Glu Ser Arg	Lys		
	65		70		75
Arg Gly Met Leu	Glu Asn Cys Ile Leu	Leu Ser Leu Phe Ala	Lys		
	80		85		90
Glu His Leu Gln	His Met Thr Glu Lys	Gln Leu Asn Leu Tyr	Asp		
	95		100		105
Arg Leu Ile Asn	Glu Pro Ser Asn Asp	Trp Asp Ile Tyr Tyr	Trp		
	110		115		120
Ala Thr Gly Arg	Arg Phe Tyr Thr Arg	Lys Trp His Ile Leu	Lys		
	125		130		135
Trp Ser Ser Thr	Asp Ser Asn Ser Ser	Gln Pro Cys Gly Gly	Gly		
	140		145		150
Arg Arg Leu Gly	Pro Glu Pro Trp Lys	Gln Gly Leu Ala Arg	Ala		
	155		160		165
Ala Ser Asp Pro	Pro Leu Leu Ala Arg	Pro Pro Gly Ala Leu	Pro		
	170		175		180
His Ser Ile Met	Met Gly Lys Leu Pro	Leu Gly Val Val Ser	Pro		
	185		190		195
Tyr Val Lys Met	Ser Ser Gly Gly Tyr	Thr Asp Pro Leu Lys	Phe		
	200		205		210
Tyr Ala Thr Ser	Tyr Cys Thr Ala Tyr	Gly Arg Glu Asp Phe	Lys		
	215		220		225
Pro Arg Val Gly	Ser His Val Gly Thr	Gly Tyr Lys Ser Asn	Phe		
	230		235		240
Gln Pro Val Val	Ser Cys Gln Ala Ser	Leu Glu Ala Leu Asp	Asn		
	245		250		255
Pro Ala Arg Gly	Glu Gln Ala Gln Asp	His Phe Gln Ser Val	Ala		
	260		265		270
Ser Gln Ser Tyr	Arg Pro Leu Glu Val	Pro Asp Gly Lys His	Pro		
	275		280		285
Leu Pro Trp Ser	Met Arg Gln Thr Ser	Ser Gly Tyr Gly Arg	Glu		
	290		295		300
Lys Pro Ser Ala	Gly Pro Pro Thr Lys	Glu Val Arg Lys Val	His		
	305		310		315
Phe Asp Thr Gln	Glu His Gly Pro Gln	Ala Ile Thr Gly Leu	Glu		
	320		325		330
Pro Arg Glu Val	Pro Leu Leu His Gln	Gln Gln Gly Gln Asp	Pro		
	335		340		345
Leu Glu Arg Glu	Asn Phe Arg His Gly	Pro Arg Phe Met Thr	Ser		
	350		355		360
Glu Tyr Asn Ser	Lys Tyr Leu Arg Asp	Pro Leu Asp Gln Pro	Asp		
	365		370		375
Phe Leu Gln Lys	Lys Ser Ile Gly Ala	Lys Glu Gly Ser Gly	Phe		
	380		385		390
Thr Lys Gln Ser	His Gln Ser Pro Ile	Val Phe Gln Pro Pro	Ser		
	395		400		405
Gln Ala Leu Pro	Gly Asp Pro Ala Leu	Leu Pro Gly Gln Ser	Val		
	410		415		420
Thr Lys Ser Asp	Phe Leu Pro Lys Thr	His Leu His Gly Asp	Glu		
	425		430		435
Phe Leu Pro Val	Leu Ala Arg Gly Ser	Lys Arg Glu Thr Ala	Phe		
	440		445		450
Ser Arg Gly Asn	Glu Arg Ile Leu Asn	Pro Arg Val Pro Pro	Pro		
	455		460		465
Cys Pro Glu Pro	Ser Ser Val Ser His	Gln Gln Phe Gln Pro	Leu		
	470		475		480
His Arg Met Gln	Gln Thr Asn Val Ala	Leu Leu Gly Arg Glu	Thr		
	485		490		495
Val Gly Lys Lys	Glu Pro Thr Gly Phe	Ser Leu Asn Asn Pro	Met		
	500		505		510
Tyr Val Arg Ser	Pro Cys Asp Pro Asp	Arg Asp Gln Arg Tyr	Leu		
	515		520		525

```

Thr Thr Tyr Asn Gln Gly Tyr Phe Glu Asn Ile Pro Lys Gly Leu
530 535 540
Asp Gln Glu Gly Trp Thr Arg Gly Gly Ile Gln Pro Gln Met Pro
545 550 555
Gly Gly Tyr Ala Leu Ser Gln Pro Val Ser Cys Met Glu Ala Thr
560 565 570
Pro Asn Pro Met Glu Ser Leu Arg His Leu His Pro His Val Gly
575 580 585
Arg Thr Leu Thr Ser Ala Asp Pro Phe Tyr Gln Asn Thr Pro His
590 595 600
Ser Ser Arg Cys Val Ala His Ser
605

```

```

<210> 9
<211> 424
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 72263451CD1

```

```

<400> 9
Met Gly Ser Ser Ser Leu Ser Glu Asp Tyr Arg Gln Cys Leu Glu
1 5 10 15
Arg Glu Leu Arg Arg Gly Arg Ala Gly Val Cys Gly Asp Pro Ser
20 25 30
Leu Arg Ala Val Leu Trp Gln Ile Leu Val Glu Asp Phe Asp Leu
35 40 45
His Gly Ala Leu Gln Asp Asp Ala Leu Ala Leu Leu Thr Asp Gly
50 55 60
Leu Trp Gly Arg Ala Asp Leu Ala Pro Ala Leu Arg Gly Leu Ala
65 70 75
Arg Ala Phe Glu Leu Leu Glu Leu Ala Ala Val His Leu Tyr Leu
80 85 90
Leu Pro Trp Arg Lys Glu Phe Thr Thr Ile Lys Thr Phe Ser Gly
95 100 105
Gly Tyr Val His Val Leu Lys Gly Val Leu Ser Asp Asp Leu Leu
110 115 120
Leu Lys Ser Phe Gln Lys Met Gly Tyr Val Arg Arg Asp Ser His
125 130 135
Arg Leu Met Val Thr Ala Leu Pro Pro Ala Cys Gln Leu Val Gln
140 145 150
Val Ala Leu Gly Cys Phe Ala Leu Arg Leu Glu Cys Glu Ile Leu
155 160 165
Gly Glu Val Leu Ala Gln Leu Gly Thr Ser Val Leu Pro Ala Glu
170 175 180
Glu Leu Leu Gln Ala Arg Arg Ala Ser Gly Asp Val Ala Ser Cys
185 190 195
Val Ala Trp Leu Gln Gln Arg Leu Ala Gln Asp Glu Glu Pro Pro
200 205 210
Pro Leu Pro Pro Arg Gly Ser Pro Ala Ala Tyr Arg Ala Pro Leu
215 220 225
Asp Leu Tyr Arg Asp Leu Gln Glu Asp Glu Gly Ser Glu Asp Ala
230 235 240
Ser Leu Tyr Gly Glu Pro Ser Pro Gly Pro Asp Ser Pro Pro Ala
245 250 255
Glu Leu Ala Tyr Arg Pro Pro Leu Trp Glu Gln Ser Ala Lys Leu
260 265 270
Trp Gly Thr Gly Gly Arg Ala Trp Glu Pro Pro Ala Glu Glu Leu
275 280 285
Pro Gln Ala Ser Ser Pro Pro Tyr Gly Ala Leu Glu Glu Gly Leu
290 295 300

```

Glu	Pro	Glu	Pro	Ser	Ala	Phe	Ser	Phe	Leu	Ser	Leu	Arg	Arg	Glu	
				305					310					315	
Leu	Ser	Arg	Pro	Gly	Asp	Leu	Ala	Thr	Pro	Glu	Ser	Ser	Ala	Ala	
				320					325					330	
Ala	Ser	Pro	Arg	Arg	Ile	Arg	Ala	Glu	Gly	Val	Pro	Ala	Ser	Ala	
				335					340					345	
Tyr	Arg	Ser	Val	Ser	Glu	Pro	Pro	Gly	Tyr	Gln	Ala	His	Ser	Cys	
				350					355					360	
Leu	Ser	Pro	Gly	Ala	Leu	Pro	Thr	Leu	Cys	Cys	Asp	Thr	Cys	Arg	
				365					370					375	
Gln	Leu	His	Ala	Ala	His	Cys	Ala	Ala	Leu	Pro	Ala	Cys	Arg	Pro	
				380					385					390	
Gly	His	Ser	Leu	Arg	Val	Leu	Leu	Gly	Asp	Ala	Gln	Arg	Arg	Leu	
				395					400					405	
Trp	Leu	Gln	Arg	Ala	Gln	Met	Asp	Thr	Leu	Leu	Tyr	Asn	Ser	Pro	
				410					415					420	
Gly	Ala	Arg	Pro												

&lt;210&gt; 10

&lt;211&gt; 913

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7502640CD1

&lt;400&gt; 10

Met	Asn	Glu	Ala	Met	Ala	Thr	Asp	Ser	Pro	Arg	Arg	Pro	Ser	Arg	
1				5					10					15	
Cys	Thr	Gly	Gly	Val	Val	Val	Arg	Pro	Gln	Ala	Val	Thr	Glu	Gln	
				20					25					30	
Ser	Tyr	Met	Glu	Ser	Val	Val	Thr	Phe	Leu	Gln	Asp	Val	Val	Pro	
				35					40					45	
Gln	Ala	Tyr	Ser	Gly	Thr	Pro	Leu	Thr	Glu	Glu	Lys	Glu	Lys	Ile	
				50					55					60	
Val	Trp	Val	Arg	Phe	Glu	Asn	Ala	Asp	Leu	Asn	Asp	Thr	Ser	Arg	
				65					70					75	
Asn	Leu	Glu	Phe	His	Glu	Ile	His	Ser	Thr	Gly	Ser	Glu	Pro	Pro	
				80					85					90	
Leu	Leu	Ile	Met	Ile	Gly	Tyr	Ser	Asp	Gly	Met	Gln	Val	Trp	Ser	
				95					100					105	
Ile	Pro	Ile	Ser	Gly	Glu	Ala	Gln	Glu	Leu	Phe	Ser	Val	Arg	His	
				110					115					120	
Gly	Pro	Ile	Arg	Ala	Ala	Arg	Ile	Leu	Pro	Ala	Pro	Gln	Phe	Gly	
				125					130					135	
Ala	Gln	Lys	Cys	Asp	Asn	Phe	Ala	Glu	Lys	Arg	Pro	Leu	Leu	Gly	
				140					145					150	
Val	Cys	Lys	Ser	Ile	Gly	Ser	Ser	Gly	Thr	Ser	Pro	Pro	Tyr	Cys	
				155					160					165	
Cys	Val	Asp	Leu	Tyr	Ser	Leu	Arg	Thr	Gly	Glu	Met	Val	Lys	Ser	
				170					175					180	
Ile	Gln	Phe	Lys	Thr	Pro	Ile	Tyr	Asp	Leu	His	Cys	Asn	Lys	Arg	
				185					190					195	
Ile	Leu	Val	Val	Val	Leu	Gln	Glu	Lys	Ile	Ala	Ala	Phe	Asp	Ser	
				200					205					210	
Cys	Thr	Phe	Thr	Lys	Lys	Phe	Phe	Val	Thr	Ser	Cys	Tyr	Pro	Cys	
				215					220					225	
Pro	Gly	Pro	Asn	Met	Asn	Pro	Ile	Ala	Leu	Gly	Ser	Arg	Trp	Leu	
				230					235					240	
Ala	Tyr	Ala	Glu	Asn	Lys	Leu	Ile	Arg	Cys	His	Gln	Ser	Arg	Gly	
				245					250					255	

Gly	Ala	Cys	Gly	Asp	Asn	Ile	Gln	Ser	Tyr	Thr	Ala	Thr	Val	Ile
				260					265					270
Ser	Ala	Ala	Lys	Thr	Leu	Lys	Ser	Gly	Leu	Thr	Met	Val	Gly	Lys
				275					280					285
Val	Val	Thr	Gln	Leu	Thr	Gly	Thr	Leu	Pro	Ser	Gly	Val	Thr	Glu
				290					295					300
Asp	Asp	Val	Ala	Ile	His	Ser	Asn	Ser	Arg	Arg	Ser	Pro	Leu	Val
				305					310					315
Pro	Gly	Ile	Ile	Thr	Val	Ile	Asp	Thr	Glu	Thr	Val	Gly	Glu	Gly
				320					325					330
Gln	Val	Leu	Val	Ser	Glu	Asp	Ser	Asp	Ser	Asp	Gly	Ile	Val	Ala
				335					340					345
His	Phe	Pro	Ala	His	Glu	Lys	Pro	Val	Cys	Cys	Met	Ala	Phe	Asn
				350					355					360
Thr	Ser	Gly	Met	Leu	Leu	Val	Thr	Thr	Asp	Thr	Leu	Gly	His	Asp
				365					370					375
Phe	His	Val	Phe	Gln	Ile	Leu	Thr	His	Pro	Trp	Ser	Ser	Ser	Gln
				380					385					390
Cys	Ala	Val	His	His	Leu	Tyr	Thr	Leu	His	Arg	Gly	Glu	Thr	Glu
				395					400					405
Ala	Lys	Val	Gln	Asp	Ile	Cys	Phe	Ser	His	Asp	Cys	Arg	Trp	Val
				410					415					420
Val	Val	Ser	Thr	Leu	Arg	Gly	Thr	Ser	His	Val	Phe	Pro	Ile	Asn
				425					430					435
Pro	Tyr	Gly	Gly	Gln	Pro	Cys	Val	Arg	Thr	His	Met	Ser	Pro	Arg
				440					445					450
Val	Val	Asn	Arg	Met	Ser	Arg	Phe	Gln	Lys	Ser	Ala	Gly	Leu	Glu
				455					460					465
Glu	Ile	Glu	Gln	Glu	Leu	Thr	Ser	Lys	Gln	Gly	Gly	Arg	Cys	Ser
				470					475					480
Pro	Val	Pro	Gly	Leu	Ser	Ser	Ser	Pro	Ser	Gly	Ser	Pro	Leu	His
				485					490					495
Gly	Lys	Leu	Asn	Ser	Gln	Asp	Ser	Tyr	Asn	Asn	Phe	Thr	Asn	Asn
				500					505					510
Asn	Pro	Gly	Asn	Pro	Arg	Leu	Ser	Pro	Leu	Pro	Ser	Leu	Met	Val
				515					520					525
Val	Met	Pro	Leu	Ala	Gln	Ile	Lys	Gln	Pro	Met	Thr	Leu	Gly	Thr
				530					535					540
Ile	Thr	Lys	Arg	Thr	Gly	Lys	Val	Lys	Pro	Pro	Pro	Gln	Ile	Ser
				545					550					555
Pro	Ser	Lys	Ser	Met	Gly	Gly	Glu	Phe	Cys	Val	Ala	Ala	Ile	Phe
				560					565					570
Gly	Thr	Ser	Arg	Ser	Trp	Phe	Ala	Asn	Asn	Ala	Gly	Leu	Lys	Arg
				575					580					585
Glu	Lys	Asp	Gln	Ser	Lys	Gln	Val	Val	Val	Glu	Ser	Leu	Tyr	Ile
				590					595					600
Ile	Ser	Cys	Tyr	Gly	Thr	Leu	Val	Glu	His	Met	Met	Glu	Pro	Arg
				605					610					615
Pro	Leu	Ser	Thr	Ala	Pro	Lys	Ile	Ser	Asp	Asp	Thr	Pro	Leu	Glu
				620					625					630
Met	Met	Thr	Ser	Pro	Arg	Ala	Ser	Trp	Thr	Leu	Val	Arg	Thr	Pro
				635					640					645
Gln	Trp	Asn	Glu	Leu	Gln	Pro	Pro	Phe	Asn	Ala	Asn	His	Pro	Leu
				650					655					660
Leu	Leu	Ala	Ala	Asp	Ala	Val	Gln	Tyr	Tyr	Gln	Phe	Leu	Leu	Ala
				665					670					675
Gly	Leu	Val	Pro	Pro	Gly	Ser	Pro	Gly	Pro	Ile	Thr	Arg	His	Gly
				680					685					690
Ser	Tyr	Asp	Ser	Leu	Ala	Ser	Asp	His	Ser	Gly	Gln	Glu	Asp	Glu
				695					700					705
Glu	Trp	Leu	Ser	Gln	Val	Glu	Ile	Val	Thr	His	Thr	Gly	Pro	His
				710					715					720
Arg	Arg	Leu	Trp	Met	Gly	Pro	Gln	Phe	Gln	Phe	Lys	Thr	Ile	His

	725		730		735
Pro Ser Gly Gln	Thr Thr Val Ile Ser	Ser Ser Ser Ser Val	Leu		
	740		745		750
Gln Ser His Gly	Pro Ser Asp Thr Pro	Gln Pro Leu Leu Asp	Phe		
	755		760		765
Asp Thr Asp Asp	Leu Asp Leu Asn Ser	Leu Arg Ile Gln Pro	Val		
	770		775		780
Arg Ser Asp Pro	Val Ser Met Pro Gly	Ser Ser Arg Pro Val	Ser		
	785		790		795
Asp Arg Arg Gly	Val Ser Thr Val Ile	Asp Ala Ala Ser Gly	Thr		
	800		805		810
Phe Asp Arg Ser	Val Thr Leu Leu Glu	Val Cys Gly Ser Trp	Pro		
	815		820		825
Glu Gly Phe Gly	Leu Arg His Met Ser	Ser Met Glu His Thr	Glu		
	830		835		840
Glu Gly Leu Arg	Glu Arg Leu Ala Asp	Ala Met Ala Glu Ser	Pro		
	845		850		855
Ser Arg Asp Val	Val Gly Ser Gly Thr	Glu Leu Gln Arg Glu	Gly		
	860		865		870
Ser Ile Glu Thr	Leu Ser Asn Ser Ser	Gly Ser Thr Ser Gly	Ser		
	875		880		885
Ile Pro Arg Asn	Phe Asp Gly Tyr Arg	Ser Pro Leu Pro Thr	Asn		
	890		895		900
Glu Ser Gln Pro	Leu Ser Leu Phe Pro	Thr Gly Phe Pro			
	905		910		

&lt;210&gt; 11

&lt;211&gt; 264

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7505807CD1

&lt;400&gt; 11

Met Cys Gly Arg Thr	Ser Cys His Leu Pro	Arg Asp Val Leu Thr		
1	5	10	15	
Arg Ala Cys Ala Tyr	Gln Asp Arg Arg Gly	Gln Gln Arg Leu Pro		
	20	25	30	
Glu Trp Arg Asp Pro	Asp Lys Tyr Cys Pro	Ser Tyr Asn Lys Ser		
	35	40	45	
Pro Gln Ser Asn Ser	Pro Val Leu Leu Ser	Arg Leu His Phe Glu		
	50	55	60	
Lys Ser Gly Ser Ile	Gly Ala Ala Asp Ser	Pro Glu Asn Trp Glu		
	65	70	75	
Lys Val Trp Asp Asn	Trp Arg Leu Leu Thr	Met Ala Gly Ile Phe		
	80	85	90	
Asp Cys Trp Glu Pro	Pro Glu Gly Gly Asp	Val Leu Tyr Ser Tyr		
	95	100	105	
Thr Ile Ile Thr Val	Asp Ser Cys Lys Gly	Leu Ser Asp Ile His		
	110	115	120	
His Arg Met Pro Ala	Ile Leu Asp Gly Glu	Glu Ala Val Ser Lys		
	125	130	135	
Trp Leu Asp Phe Gly	Glu Val Ser Thr Gln	Glu Ala Leu Lys Leu		
	140	145	150	
Ile His Pro Thr Glu	Asn Ile Thr Phe His	Ala Val Ser Ser Val		
	155	160	165	
Val Asn Asn Ser Arg	Asn Asn Thr Pro Glu	Cys Leu Ala Pro Val		
	170	175	180	
Asp Leu Val Val Lys	Lys Glu Leu Arg Ala	Ser Gly Ser Ser Gln		
	185	190	195	
Arg Met Leu Gln Trp	Leu Ala Thr Lys Ser	Pro Lys Lys Glu Asp		



	200		205		210
Ser Lys Thr Pro	Gln Lys Glu Glu Ser	Asp Val Pro Gln Trp	Ser		
	215		220		225
Ser Gln Phe Leu	Gln Lys Ser Pro Leu	Pro Thr Lys Arg Gly	Thr		
	230		235		240
Ala Gly Leu Leu	Glu Gln Trp Leu Lys	Arg Glu Lys Glu Glu	Glu		
	245		250		255
Pro Val Ala Lys	Arg Pro Tyr Ser Gln				
	260				

<210> 12  
 <211> 553  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7506413CD1

<400> 12

Met Leu Ser Leu Cys	Ala Gln Asp Leu Thr	Gln Met Leu Ala Leu	
1	5	10	15
Ser Arg His Ser	Leu Leu Ser Pro Leu	Leu Ser Val Thr Ser	Phe
	20	25	30
Arg Arg Phe Tyr	Arg Gly Asp Ser Pro	Thr Asp Ser Gln Lys	Asp
	35	40	45
Met Ile Glu Ile	Pro Leu Pro Pro Trp	Gln Glu Arg Thr Asp	Glu
	50	55	60
Ser Ile Glu Thr	Lys Arg Ala Arg Leu	Leu Tyr Glu Ser Arg	Lys
	65	70	75
Arg Gly Met Leu	Glu Asn Cys Ile Leu	Leu Ser Leu Phe Ala	Lys
	80	85	90
Glu His Leu Gln	His Met Thr Glu Lys	Gln Leu Asn Leu Tyr	Asp
	95	100	105
Arg Leu Ile Asn	Glu Pro Ser Asn Asp	Trp Asp Ile Tyr Tyr	Trp
	110	115	120
Ala Thr Gly Arg	Arg Phe Tyr Thr Arg	Lys Trp His Ile Leu	Lys
	125	130	135
Trp Ser Ser Thr	Asp Ser Asn Ser Ser	Gln Pro Cys Gly Gly	Gly
	140	145	150
Arg Arg Leu Gly	Pro Glu Pro Trp Lys	Gln Gly Leu Ala Arg	Ala
	155	160	165
Ala Ser Asp Pro	Pro Leu Leu Ala Arg	Pro Pro Gly Ala Leu	Pro
	170	175	180
His Ser Ile Met	Met Gly Lys Leu Pro	Leu Gly Val Val Ser	Pro
	185	190	195
Tyr Val Lys Met	Ser Ser Gly Gly Tyr	Thr Asp Pro Leu Lys	Phe
	200	205	210
Tyr Ala Thr Ser	Tyr Cys Thr Ala Tyr	Gly Arg Glu Asp Phe	Lys
	215	220	225
Pro Arg Val Gly	Ser His Val Gly Thr	Gly Tyr Lys Ser Asn	Phe
	230	235	240
Gln Pro Val Val	Ser Cys Gln Ala Ser	Leu Glu Ala Leu Asp	Asn
	245	250	255
Pro Ala Arg Gly	Glu Gln Ala Gln Asp	His Phe Gln Ser Val	Ala
	260	265	270
Ser Gln Ser Tyr	Arg Pro Leu Glu Val	Pro Asp Gly Lys His	Pro
	275	280	285
Leu Pro Trp Ser	Met Arg Gln Thr Ser	Ser Gly Tyr Gly Arg	Glu
	290	295	300
Lys Pro Ser Ala	Gly Pro Pro Thr Lys	Glu Val Arg Lys Val	His
	305	310	315
Phe Asp Thr Gln	Glu His Gly Pro Gln	Ala Ile Thr Gly Leu	Glu

Pro Arg Glu Val	320	Pro Leu Leu His Gln	325	Gln Gln Gly Gln Asp	330
	335		340		345
Leu Glu Arg Glu	350	Asn Phe Arg His Gly	355	Pro Arg Phe Met Thr	360
	365		370		375
Glu Tyr Asn Ser	380	Lys Tyr Leu Arg Asp	385	Pro Leu Asp Gln Pro	390
	395		400		405
Phe Leu Gln Lys	410	Lys Ser Ile Gly Ala	415	Glu Gly Ser Gly Phe	420
	425		430		435
Thr Lys Gln Ser	440	His Gln Ser Pro Ile	445	Val Phe Gln Pro Pro	450
	455		460		465
Gln Ala Leu Pro	470	Gly Asp Pro Ala Leu	475	Leu Pro Gly Gln Ser	480
	485		490		495
Thr Lys Ser Asp	500	Phe Leu Pro Lys Thr	505	His Leu His Gly Asp	510
	515		520		525
Phe Leu Pro Val	530	Leu Ala Arg Gly Ser	535	Lys Arg Glu Thr Ala	540
	545		550		
Ser Arg Gly Asn		Glu Arg Ile Leu Asn		Pro Arg Val Pro Pro	
Cys Pro Glu Pro		Ser Ser Val Ser His		Gln Gln Phe Gln Pro	
His Arg Met Gln		Gln Thr Asn Val Ala		Leu Leu Gly Arg Glu	
Val Gly Lys Lys		Glu Pro Thr Gly Phe		Ser Leu Asn Asn Pro	
Tyr Val Arg Ser		Pro Cys Asp Pro Asp		Arg Asp Gln Arg Tyr	
Thr Thr Tyr Asn		Gln Gly Thr Leu Thr		Ser Ala Asp Pro Phe	
Gln Asn Thr Pro		His Ser Ser Arg Cys		Val Ala His Ser	

&lt;210&gt; 13

&lt;211&gt; 263

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1283631CD1

&lt;400&gt; 13

Met Gly Lys Leu Thr	1	Thr Met Pro Ala Gly	10	Leu Ile Tyr Ala Ser	15
Val Ser Val His Ala	20	Ala Lys Gln Glu Glu	25	Ser Lys Lys Gln Leu	30
Val Lys Pro Glu Gln	35	Leu Pro Ile Tyr Thr	40	Ala Pro Pro Leu Gln	45
Ser Lys Tyr Val Glu	50	Glu Gln Pro Gly His	55	Leu Gln Met Gly Phe	60
Ala Ser Ile Arg Thr	65	Ala Thr Gly Cys Tyr	70	Ile Gly Trp Cys Lys	75
Gly Val Tyr Val Phe	80	Val Lys Asn Gly Ile	85	Met Asp Thr Val Gln	90
Phe Gly Lys Asp Ala	95	Tyr Val Tyr Leu Lys	100	Asn Pro Pro Arg Asp	105
Phe Leu Pro Lys Met	110	Gly Val Ile Thr Val	115	Ser Gly Leu Ala Gly	120
Leu Val Ser Ala Arg	125	Lys Gly Ser Lys Phe	130	Lys Lys Ile Thr Tyr	135
Pro Leu Gly Leu Ala	140	Thr Leu Gly Ala Thr	145	Val Cys Tyr Pro Val	150
Gln Ser Val Ile Ile		Ala Lys Val Thr Ala		Lys Lys Val Tyr Ala	

	155		160		165
Thr Ser Gln Gln	Ile Phe Gly Ala Val	Lys Ser Leu Trp Thr	Lys		
	170		175		180
Ser Ser Lys Glu	Glu Ser Leu Pro Lys	Pro Lys Glu Lys Thr	Lys		
	185		190		195
Leu Gly Ser Ser	Ser Glu Ile Glu Val	Pro Ala Lys Thr Thr	His		
	200		205		210
Val Leu Lys His	Ser Val Pro Leu Pro	Thr Glu Leu Ser Ser	Glu		
	215		220		225
Ala Lys Thr Lys	Ser Glu Ser Thr Ser	Gly Ala Thr Gln Phe	Met		
	230		235		240
Pro Asp Pro Lys	Leu Met Asp His Gly	Gln Ser His Pro Glu	Asp		
	245		250		255
Ile Asp Met Tyr	Ser Thr Arg Ser				
	260				

<210> 14  
 <211> 1449  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1740413CD1

<400> 14

Met Glu Val Leu Arg Asp Pro Ile Lys Lys Asn Ser Ser Glu Ser		
1 5 10 15		
Lys Pro Ala Gln Ser Gly Phe Ser Arg Gly Asn Ser Pro Leu Ser		
20 25 30		
Cys Pro Glu Ser Val Glu Ala Ser Pro Ala Val Asn Glu Lys Ser		
35 40 45		
Val Tyr Ser Thr His Asn Tyr Gly Thr Thr Gln Arg His Gly Cys		
50 55 60		
Arg Gly Leu Pro Tyr Ala Thr Ile Ile Pro Arg Ser Asp Leu Asn		
65 70 75		
Gly Leu Pro Ser Pro Val Glu Glu Arg Cys Gly Asp Ser Pro Asn		
80 85 90		
Ser Glu Gly Glu Thr Val Pro Thr Trp Cys Pro Cys Gly Leu Ser		
95 100 105		
Gln Asp Gly Phe Leu Leu Asn Cys Asp Lys Cys Arg Gly Met Ser		
110 115 120		
Arg Gly Lys Val Ile Arg Leu His Arg Arg Lys Gln Asp Asn Ile		
125 130 135		
Ser Gly Gly Asp Ser Ser Ala Thr Glu Ser Trp Asp Glu Glu Leu		
140 145 150		
Ser Pro Ser Thr Val Leu Tyr Thr Ala Thr Gln His Thr Pro Thr		
155 160 165		
Ser Ile Thr Leu Thr Val Arg Arg Thr Lys Pro Lys Lys Arg Lys		
170 175 180		
Lys Ser Pro Glu Lys Gly Arg Ala Ala Pro Lys Thr Lys Lys Ile		
185 190 195		
Lys Asn Ser Pro Ser Glu Ala Gln Asn Leu Asp Glu Asn Thr Thr		
200 205 210		
Glu Gly Trp Glu Asn Arg Ile Arg Leu Trp Thr Asp Gln Tyr Glu		
215 220 225		
Glu Ala Phe Thr Asn Gln Tyr Ser Ala Asp Val Gln Asn Ala Leu		
230 235 240		
Glu Gln His Leu His Ser Ser Lys Glu Phe Val Gly Lys Pro Thr		
245 250 255		
Ile Leu Asp Thr Ile Asn Lys Thr Glu Leu Ala Cys Asn Asn Thr		
260 265 270		
Val Ile Gly Ser Gln Met Gln Leu Gln Leu Gly Arg Val Thr Arg		

	275		280		285
Val Gln Lys His	Arg Lys Ile Leu Arg	Ala Ala Arg Asp Leu	Ala		
	290		295		300
Leu Asp Thr Leu	Ile Ile Glu Tyr Arg	Gly Lys Val Met Leu	Arg		
	305		310		315
Gln Gln Phe Glu	Val Asn Gly His Phe	Phe Lys Lys Pro Tyr	Pro		
	320		325		330
Phe Val Leu Phe	Tyr Ser Lys Phe Asn	Gly Val Glu Met Cys	Val		
	335		340		345
Asp Ala Arg Thr	Phe Gly Asn Asp Ala	Arg Phe Ile Arg Arg	Ser		
	350		355		360
Cys Thr Pro Asn	Ala Glu Val Arg His	Met Ile Ala Asp Gly	Met		
	365		370		375
Ile His Leu Cys	Ile Tyr Ala Val Ser	Ala Ile Thr Lys Asp	Ala		
	380		385		390
Glu Val Thr Ile	Ala Phe Asp Tyr Glu	Tyr Ser Asn Cys Asn	Tyr		
	395		400		405
Lys Val Asp Cys	Ala Cys His Lys Gly	Asn Arg Asn Cys Pro	Ile		
	410		415		420
Gln Lys Arg Asn	Pro Asn Ala Thr Glu	Leu Pro Leu Leu Pro	Pro		
	425		430		435
Pro Pro Ser Leu	Pro Thr Ile Gly Ala	Glu Thr Arg Arg Arg	Lys		
	440		445		450
Ala Arg Arg Lys	Glu Leu Glu Met Glu	Gln Gln Asn Glu Ala	Ser		
	455		460		465
Glu Glu Asn Asn	Asp Gln Gln Ser Gln	Glu Val Pro Glu Lys	Val		
	470		475		480
Thr Val Ser Ser	Asp His Glu Glu Val	Asp Asn Pro Glu Glu	Lys		
	485		490		495
Pro Glu Glu Glu	Lys Glu Glu Val Ile	Asp Asp Gln Glu Asn	Leu		
	500		505		510
Ala His Ser Arg	Arg Thr Arg Glu Asp	Arg Lys Val Glu Ala	Ile		
	515		520		525
Met His Ala Phe	Glu Asn Leu Glu Lys	Arg Lys Lys Arg Arg	Asp		
	530		535		540
Gln Pro Leu Glu	Gln Ser Asn Ser Asp	Val Glu Ile Thr Thr	Thr		
	545		550		555
Thr Ser Glu Thr	Pro Val Gly Glu Glu	Thr Lys Thr Glu Ala	Pro		
	560		565		570
Glu Ser Glu Val	Ser Asn Ser Val Ser	Asn Val Thr Ile Pro	Ser		
	575		580		585
Thr Pro Gln Ser	Val Gly Val Asn Thr	Arg Arg Ser Ser Gln	Ala		
	590		595		600
Gly Asp Ile Ala	Ala Glu Lys Leu Val	Pro Lys Pro Pro Pro	Ala		
	605		610		615
Lys Pro Ser Arg	Pro Arg Pro Lys Ser	Arg Ile Ser Arg Tyr	Arg		
	620		625		630
Thr Ser Ser Ala	Gln Arg Leu Lys Arg	Gln Lys Gln Ala Asn	Ala		
	635		640		645
Gln Gln Ala Glu	Leu Ser Gln Ala Ala	Leu Glu Glu Gly Gly	Ser		
	650		655		660
Asn Ser Leu Val	Thr Pro Thr Glu Ala	Gly Ser Leu Asp Ser	Ser		
	665		670		675
Gly Glu Asn Arg	Pro Leu Thr Gly Ser	Asp Pro Thr Val Val	Ser		
	680		685		690
Ile Thr Gly Ser	His Val Asn Arg Ala	Ala Ser Lys Tyr Pro	Lys		
	695		700		705
Thr Lys Lys Tyr	Leu Val Thr Glu Trp	Leu Asn Asp Lys Ala	Glu		
	710		715		720
Lys Gln Glu Cys	Pro Val Glu Cys Pro	Leu Arg Ile Thr Thr	Asp		
	725		730		735
Pro Thr Val Leu	Ala Thr Thr Leu Asn	Met Leu Pro Gly Leu	Ile		
	740		745		750

His	Ser	Pro	Leu	Ile	Cys	Thr	Thr	Pro	Lys	His	Tyr	Ile	Arg	Phe
				755					760					765
Gly	Ser	Pro	Phe	Ile	Pro	Glu	Arg	Arg	Arg	Arg	Pro	Leu	Leu	Pro
				770					775					780
Asp	Gly	Thr	Phe	Ser	Ser	Cys	Lys	Lys	Arg	Trp	Ile	Lys	Gln	Ala
				785					790					795
Leu	Glu	Glu	Gly	Met	Thr	Gln	Thr	Ser	Ser	Val	Pro	Gln	Glu	Thr
				800					805					810
Arg	Thr	Gln	His	Leu	Tyr	Gln	Ser	Asn	Glu	Asn	Ser	Ser	Ser	Ser
				815					820					825
Ser	Ile	Cys	Lys	Asp	Asn	Ala	Asp	Leu	Leu	Ser	Pro	Leu	Lys	Lys
				830					835					840
Trp	Lys	Ser	Arg	Tyr	Leu	Met	Glu	Gln	Asn	Val	Thr	Lys	Leu	Leu
				845					850					855
Arg	Pro	Leu	Ser	Pro	Val	Thr	Pro	Pro	Pro	Pro	Asn	Ser	Gly	Ser
				860					865					870
Lys	Ser	Pro	Gln	Leu	Ala	Thr	Pro	Gly	Ser	Ser	His	Pro	Gly	Glu
				875					880					885
Glu	Glu	Cys	Arg	Asn	Gly	Tyr	Ser	Leu	Met	Phe	Ser	Pro	Val	Thr
				890					895					900
Ser	Leu	Thr	Thr	Ala	Ser	Arg	Cys	Asn	Thr	Pro	Leu	Gln	Phe	Glu
				905					910					915
Leu	Cys	His	Arg	Lys	Asp	Leu	Asp	Leu	Ala	Lys	Val	Gly	Tyr	Leu
				920					925					930
Asp	Ser	Asn	Thr	Asn	Ser	Cys	Ala	Asp	Arg	Pro	Ser	Leu	Leu	Asn
				935					940					945
Ser	Gly	His	Ser	Asp	Leu	Ala	Pro	His	Pro	Ser	Leu	Gly	Pro	Thr
				950					955					960
Ser	Glu	Thr	Gly	Phe	Pro	Ser	Arg	Ser	Gly	Asp	Gly	His	Gln	Thr
				965					970					975
Leu	Val	Arg	Asn	Ser	Asp	Gln	Ala	Phe	Arg	Thr	Glu	Phe	Asn	Leu
				980					985					990
Met	Tyr	Ala	Tyr	Ser	Pro	Leu	Asn	Ala	Met	Pro	Arg	Ala	Asp	Gly
				995					1000					1005
Leu	Tyr	Arg	Gly	Ser	Pro	Leu	Val	Gly	Asp	Arg	Lys	Pro	Leu	His
				1010					1015					1020
Leu	Asp	Gly	Gly	Tyr	Cys	Ser	Pro	Ala	Glu	Gly	Phe	Ser	Ser	Arg
				1025					1030					1035
Tyr	Glu	His	Gly	Leu	Met	Lys	Asp	Leu	Ser	Arg	Gly	Ser	Leu	Ser
				1040					1045					1050
Pro	Gly	Gly	Glu	Arg	Ala	Cys	Glu	Gly	Val	Pro	Ser	Ala	Pro	Gln
				1055					1060					1065
Asn	Pro	Pro	Gln	Arg	Lys	Lys	Val	Ser	Leu	Leu	Glu	Tyr	Arg	Lys
				1070					1075					1080
Arg	Lys	Gln	Glu	Ala	Lys	Glu	Asn	Ser	Ala	Gly	Gly	Gly	Gly	Asp
				1085					1090					1095
Ser	Ala	Gln	Ser	Lys	Ser	Lys	Ser	Ala	Gly	Ala	Gly	Gln	Gly	Ser
				1100					1105					1110
Ser	Asn	Ser	Val	Ser	Asp	Thr	Gly	Ala	His	Gly	Val	Gln	Gly	Ser
				1115					1120					1125
Ser	Ala	Arg	Thr	Pro	Ser	Ser	Pro	His	Lys	Lys	Phe	Ser	Pro	Ser
				1130					1135					1140
His	Ser	Ser	Met	Ser	His	Leu	Glu	Ala	Val	Ser	Pro	Ser	Asp	Ser
				1145					1150					1155
Arg	Gly	Thr	Ser	Ser	Ser	His	Cys	Arg	Pro	Gln	Glu	Asn	Ile	Ser
				1160					1165					1170
Ser	Arg	Trp	Met	Val	Pro	Thr	Ser	Val	Glu	Arg	Leu	Arg	Glu	Gly
				1175					1180					1185
Gly	Ser	Ile	Pro	Lys	Val	Leu	Arg	Ser	Ser	Val	Arg	Val	Ala	Gln
				1190					1195					1200
Lys	Gly	Glu	Pro	Ser	Pro	Thr	Trp	Glu	Ser	Asn	Ile	Thr	Glu	Lys
				1205					1210					1215
Asp	Ser	Asp	Pro	Ala	Asp	Gly	Glu	Gly	Pro	Glu	Thr	Leu	Ser	Ser

1220	1225	1230
Ala Leu Ser Lys Gly	Ala Thr Val Tyr Ser	Pro Ser Arg Tyr Ser
1235	1240	1245
Tyr Gln Leu Leu Gln	Cys Asp Ser Pro Arg	Thr Glu Ser Gln Ser
1250	1255	1260
Leu Leu Gln Gln Ser	Ser Ser Pro Phe Arg	Gly His Pro Thr Gln
1265	1270	1275
Ser Pro Gly Tyr Ser	Tyr Arg Thr Thr Ala	Leu Arg Pro Gly Asn
1280	1285	1290
Pro Pro Ser His Gly	Ser Ser Glu Ser Ser	Leu Ser Ser Thr Ser
1295	1300	1305
Tyr Ser Ser Pro Ala	His Pro Val Ser Thr	Asp Ser Leu Ala Pro
1310	1315	1320
Phe Thr Gly Thr Pro	Gly Tyr Phe Ser Ser	Gln Pro His Ser Gly
1325	1330	1335
Asn Ser Thr Gly Ser	Asn Leu Pro Arg Arg	Ser Cys Pro Ser Ser
1340	1345	1350
Ala Ala Ser Pro Thr	Leu Gln Gly Pro Ser	Asp Ser Pro Thr Ser
1355	1360	1365
Asp Ser Val Ser Gln	Ser Ser Thr Gly Thr	Leu Ser Ser Thr Ser
1370	1375	1380
Phe Pro Gln Asn Ser	Arg Ser Ser Leu Pro	Ser Asp Leu Arg Thr
1385	1390	1395
Ile Ser Leu Pro Ser	Ala Gly Gln Ser Ala	Val Tyr Gln Ala Ser
1400	1405	1410
Arg Val Ser Ala Val	Ser Asn Ser Gln His	Tyr Pro His Arg Gly
1415	1420	1425
Ser Gly Gly Val His	Gln Tyr Arg Leu Gln	Pro Leu Gln Gly Ser
1430	1435	1440
Gly Val Lys Thr Gln	Thr Gly Leu Ser	
1445		

&lt;210&gt; 15

&lt;211&gt; 400

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1951731CD1

&lt;400&gt; 15

Met Val Phe His Ala	Pro Ser Arg Ala	Leu Leu Val His Gly Gly
1	5	10 15
His Arg Pro Ser Thr	Ala Arg Phe Ser	Val Arg Val Asn Ser Thr
20	25	30
Glu Leu Phe His Val	Asp Arg His Val	Trp Thr Thr Leu Lys Gly
35	40	45
Arg Asp Gly Leu Gln	Gly Pro Arg Glu	Arg Ala Phe His Thr Ala
50	55	60
Ser Val Leu Gly Asn	Tyr Met Val Val	Tyr Gly Gly Asn Val His
65	70	75
Thr His Tyr Gln Glu	Lys Cys Tyr Glu	Asp Gly Ile Phe Phe
80	85	90
Tyr His Leu Gly Cys	His Gln Trp Val	Ser Gly Ala Glu Leu Ala
95	100	105
Pro Pro Gly Thr Pro	Glu Gly Arg Ala	Ala Pro Pro Ser Gly Arg
110	115	120
Tyr Ser His Val Ala	Ala Val Leu Gly	Gly Ser Val Leu Leu Val
125	130	135
Ala Gly Gly Tyr Ser	Gly Arg Pro Arg	Gly Asp Leu Met Ala Tyr
140	145	150
Lys Val Pro Pro Phe	Val Phe Gln Ala	Pro Ala Pro Asp Tyr His

155	160	165
Leu Asp Tyr Cys Ser Met Tyr Thr Asp His Ser Val Cys Ser Arg		
170	175	180
Asp Pro Glu Cys Ser Trp Cys Gln Gly Ala Cys Gln Ala Ala Pro		
185	190	195
Pro Pro Gly Thr Pro Leu Gly Ala Cys Pro Ala Ala Ser Cys Leu		
200	205	210
Gly Leu Gly Arg Leu Leu Gly Asp Cys Gln Ala Cys Leu Ala Phe		
215	220	225
Ser Ser Pro Thr Ala Pro Pro Arg Gly Pro Gly Thr Leu Gly Trp		
230	235	240
Cys Val His Asn Glu Ser Cys Leu Pro Arg Pro Glu Gln Ala Arg		
245	250	255
Cys Arg Gly Glu Gln Ile Ser Gly Thr Val Gly Trp Trp Gly Pro		
260	265	270
Ala Pro Val Phe Val Thr Ser Leu Glu Ala Cys Val Thr Gln Ser		
275	280	285
Phe Leu Pro Gly Leu His Leu Leu Thr Phe Gln Gln Pro Pro Asn		
290	295	300
Thr Ser Gln Pro Asp Lys Val Ser Ile Val Arg Ser Thr Thr Ile		
305	310	315
Thr Leu Thr Pro Ser Ala Glu Thr Asp Val Ser Leu Val Tyr Arg		
320	325	330
Gly Phe Ile Tyr Pro Met Leu Pro Gly Gly Pro Gly Gly Pro Gly		
335	340	345
Ala Glu Asp Val Ala Val Trp Thr Arg Ala Gln Arg Leu His Val		
350	355	360
Leu Ala Arg Met Ala Arg Gly Pro Asp Thr Glu Asn Met Val Arg		
365	370	375
Pro Pro Gly Thr Phe Arg Gly Leu Phe Met Val Asn Arg Glu Leu		
380	385	390
Gly Val Trp Tyr Arg Ala Ser Val Val Ile		
395	400	

&lt;210&gt; 16

&lt;211&gt; 226

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3741930CD1

&lt;400&gt; 16

Met Ser Arg Ile Tyr His Asp Gly Ala Leu Arg Asn Lys Ala Val		
1	5	10
Gln Ser Val Arg Leu Pro Gly Ala Trp Asp Pro Ala Ala His Gln		
20	25	30
Gly Gly Asn Gly Val Leu Leu Glu Gly Glu Leu Ile Asp Val Ser		
35	40	45
Arg His Ser Ile Leu Asp Thr His Gly Arg Lys Glu Arg Tyr Tyr		
50	55	60
Val Leu Tyr Ile Arg Pro Ser His Ile His Arg Arg Lys Phe Asp		
65	70	75
Ala Lys Gly Asn Glu Ile Glu Pro Asn Phe Ser Ala Thr Arg Lys		
80	85	90
Val Asn Thr Gly Phe Leu Met Ser Ser Tyr Lys Val Glu Ala Lys		
95	100	105
Gly Asp Thr Asp Arg Leu Thr Pro Glu Ala Leu Lys Gly Leu Val		
110	115	120
Asn Lys Pro Glu Leu Leu Ala Leu Thr Glu Ser Leu Thr Pro Asp		
125	130	135
His Thr Val Ala Phe Trp Met Pro Glu Ser Glu Met Glu Val Met		

	140		145		150
Glu Leu Glu Leu	Gly Ala Gly Val Arg	Leu Lys Thr Arg Gly Asp			
	155		160		165
Gly Pro Phe Leu	Asp Ser Leu Ala Lys	Leu Glu Ala Gly Thr Val			
	170		175		180
Thr Lys Cys Asn	Phe Thr Gly Asp Gly	Lys Thr Gly Ala Ser Trp			
	185		190		195
Thr Asp Asn Ile	Met Ala Gln Lys Cys	Ser Lys Gly Ala Ala Ala			
	200		205		210
Glu Ile Arg Glu	Gln Gly Asp Gly Ala	Glu Asp Glu Glu Trp Asp			
	215		220		225
Asp					

<210> 17  
 <211> 715  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 5402506CD1

<400> 17

Met Asp Leu Ile	Leu Asn Arg Met Asp Tyr	Leu Gln Val Gly Val	
1	5	10	15
Thr Ser Gln Lys	Thr Met Lys Leu Ile Pro	Ala Ser Arg His Arg	
	20	25	30
Ala Thr Gln Lys	Val Val Ile Gly Asp His	Asp Gly Val Val Met	
	35	40	45
Cys Phe Gly Met	Lys Lys Gly Glu Ala Ala	Val Phe Lys Thr	
	50	55	60
Leu Pro Gly Pro	Lys Ile Ala Arg Leu Glu	Leu Gly Gly Val Ile	
	65	70	75
Asn Thr Pro Gln	Glu Lys Ile Phe Ile Ala	Ala Ala Ser Glu Ile	
	80	85	90
Arg Gly Phe Thr	Lys Arg Gly Lys Gln Phe	Leu Ser Phe Glu Thr	
	95	100	105
Asn Leu Thr Glu	Ser Ile Lys Ala Met His	Ile Ser Gly Ser Asp	
	110	115	120
Leu Phe Leu Ser	Ala Ser Tyr Ile Tyr Asn	His Tyr Cys Asp Cys	
	125	130	135
Lys Asp Gln His	Tyr Tyr Leu Ser Gly Asp	Lys Ile Asn Asp Val	
	140	145	150
Ile Cys Leu Pro	Val Glu Arg Leu Ser Arg	Ile Thr Pro Val Leu	
	155	160	165
Ala Cys Gln Asp	Arg Val Leu Arg Val Leu	Gln Gly Ser Asp Val	
	170	175	180
Met Tyr Ala Val	Glu Val Pro Gly Pro Pro	Thr Val Leu Ala Leu	
	185	190	195
His Asn Gly Asn	Gly Gly Asp Ser Gly Glu	Asp Leu Leu Phe Gly	
	200	205	210
Thr Ser Asp Gly	Lys Leu Ala Leu Ile Gln	Ile Thr Thr Ser Lys	
	215	220	225
Pro Val Arg Lys	Trp Glu Ile Gln Asn Glu	Lys Lys Arg Gly Gly	
	230	235	240
Ile Leu Cys Ile	Asp Ser Phe Asp Ile Val	Gly Asp Gly Val Lys	
	245	250	255
Asp Leu Leu Val	Gly Arg Asp Asp Gly Met	Val Glu Val Tyr Ser	
	260	265	270
Phe Asp Asn Ala	Asn Glu Pro Val Leu Arg	Phe Asp Gln Met Leu	
	275	280	285
Ser Glu Ser Val	Thr Ser Ile Gln Gly Gly	Cys Val Gly Lys Asp	



	290		295		300
Ser Tyr Asp Glu	Ile Val Val Ser Thr	Tyr Ser Gly Trp Val	Thr		
	305		310		315
Gly Leu Thr Thr	Glu Pro Ile His Lys	Glu Ser Gly Pro Gly	Glu		
	320		325		330
Glu Leu Lys Ile	Asn Gln Glu Met Gln	Asn Lys Ile Ser Ser	Leu		
	335		340		345
Arg Asn Glu Leu	Glu His Leu Gln Tyr	Lys Val Leu Gln Glu	Arg		
	350		355		360
Glu Asn Tyr Gln	Gln Ser Ser Gln Ser	Ser Lys Ala Lys Ser	Ala		
	365		370		375
Val Pro Ser Phe	Gly Ile Asn Asp Lys	Phe Thr Leu Asn Lys	Asp		
	380		385		390
Asp Ala Ser Tyr	Ser Leu Ile Leu Glu	Val Gln Thr Ala Ile	Asp		
	395		400		405
Asn Val Leu Ile	Gln Ser Asp Val Pro	Ile Asp Leu Leu Asp	Val		
	410		415		420
Asp Lys Asn Ser	Ala Val Val Ser Phe	Ser Ser Cys Asp Ser	Glu		
	425		430		435
Ser Asn Asp Asn	Phe Leu Leu Ala Thr	Tyr Arg Cys Gln Ala	Asp		
	440		445		450
Thr Thr Arg Leu	Glu Leu Lys Ile Arg	Ser Ile Glu Gly Gln	Tyr		
	455		460		465
Gly Thr Leu Gln	Ala Tyr Val Thr Pro	Arg Ile Gln Pro Lys	Thr		
	470		475		480
Cys Gln Val Arg	Gln Tyr His Ile Lys	Pro Leu Ser Leu His	Gln		
	485		490		495
Arg Thr His Phe	Ile Asp His Asp Arg	Pro Met Asn Thr Leu	Thr		
	500		505		510
Leu Thr Gly Gln	Phe Ser Phe Ala Glu	Val His Ser Trp Val	Val		
	515		520		525
Phe Cys Leu Pro	Glu Val Pro Glu Lys	Pro Pro Ala Gly Glu	Cys		
	530		535		540
Val Thr Phe Tyr	Phe Gln Asn Thr Phe	Leu Asp Thr Gln Leu	Glu		
	545		550		555
Ser Thr Tyr Arg	Lys Gly Glu Gly Val	Phe Lys Ser Asp Asn	Ile		
	560		565		570
Ser Thr Ile Ser	Ile Leu Lys Asp Val	Leu Ser Lys Glu Ala	Thr		
	575		580		585
Lys Arg Lys Ile	Asn Leu Asn Ile Ser	Tyr Glu Ile Asn Glu	Val		
	590		595		600
Ser Val Lys His	Thr Leu Lys Leu Ile	His Pro Lys Leu Glu	Tyr		
	605		610		615
Gln Leu Leu Leu	Ala Lys Lys Val Gln	Leu Ile Asp Ala Leu	Lys		
	620		625		630
Glu Leu Gln Ile	His Glu Gly Asn Thr	Asn Phe Leu Ile Pro	Glu		
	635		640		645
Tyr His Cys Ile	Leu Glu Glu Ala Asp	His Leu Gln Glu Glu	Tyr		
	650		655		660
Lys Lys Gln Pro	Ala His Leu Glu Arg	Leu Tyr Gly Met Ile	Thr		
	665		670		675
Asp Leu Phe Ile	Asp Lys Phe Lys Phe	Lys Gly Thr Asn Val	Lys		
	680		685		690
Thr Lys Val Pro	Leu Leu Leu Glu Ile	Leu Asp Ser Tyr Asp	Gln		
	695		700		705
Asn Ala Leu Ile	Ser Phe Phe Asp Ala	Ala			
	710		715		

&lt;210&gt; 18

&lt;211&gt; 364

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 71081333CD1

&lt;400&gt; 18

Met	Pro	Arg	Gly	Arg	Cys	Arg	Gln	Gln	Gly	Pro	Arg	Ile	Pro	Ile
1				5					10					15
Trp	Ala	Ala	Ala	Asn	Tyr	Ala	Asn	Ala	His	Pro	Trp	Gln	Gln	Met
				20					25					30
Asp	Lys	Ala	Ser	Pro	Gly	Val	Ala	Tyr	Thr	Pro	Leu	Val	Asp	Pro
				35					40					45
Trp	Ile	Glu	Arg	Pro	Cys	Cys	Gly	Asp	Thr	Val	Cys	Val	Arg	Thr
				50					55					60
Thr	Met	Glu	Gln	Lys	Ser	Thr	Ala	Ser	Gly	Thr	Cys	Gly	Gly	Lys
				65					70					75
Pro	Ala	Glu	Arg	Gly	Pro	Leu	Ala	Gly	His	Met	Pro	Ser	Ser	Arg
				80					85					90
Pro	His	Arg	Val	Asp	Phe	Cys	Trp	Val	Pro	Gly	Ser	Asp	Pro	Gly
				95					100					105
Thr	Phe	Asp	Gly	Ser	Pro	Trp	Leu	Leu	Asp	Arg	Phe	Leu	Ala	Gln
				110					115					120
Leu	Gly	Asp	Tyr	Met	Ser	Phe	His	Phe	Glu	His	Tyr	Gln	Asp	Asn
				125					130					135
Ile	Ser	Arg	Val	Cys	Glu	Ile	Leu	Arg	Arg	Leu	Thr	Gly	Arg	Ala
				140					145					150
Gln	Ala	Trp	Ala	Ala	Pro	Tyr	Leu	Asp	Gly	Asp	Leu	Pro	Leu	Pro
				155					160					165
Asp	Asp	Tyr	Glu	Leu	Phe	Cys	Gln	Asp	Leu	Lys	Glu	Val	Val	Gln
				170					175					180
Asp	Pro	Asn	Ser	Phe	Ala	Glu	Tyr	His	Ala	Val	Val	Thr	Cys	Pro
				185					190					195
Leu	Pro	Leu	Ala	Ser	Ser	Gln	Leu	Pro	Val	Ala	Pro	Gln	Leu	Pro
				200					205					210
Val	Val	Arg	Gln	Tyr	Leu	Ala	Arg	Phe	Leu	Glu	Gly	Leu	Ala	Leu
				215					220					225
Asp	Met	Gly	Thr	Ala	Pro	Arg	Ser	Leu	Pro	Ala	Ala	Met	Ala	Thr
				230					235					240
Pro	Ala	Val	Ser	Gly	Ser	Asn	Ser	Val	Ser	Arg	Ser	Ala	Leu	Phe
				245					250					255
Glu	Gln	Gln	Leu	Thr	Lys	Glu	Ser	Thr	Pro	Gly	Pro	Lys	Glu	Pro
				260					265					270
Pro	Val	Leu	Pro	Ser	Ser	Thr	Cys	Ser	Ser	Lys	Pro	Gly	Pro	Val
				275					280					285
Glu	Pro	Ala	Ser	Ser	Gln	Pro	Glu	Glu	Ala	Ala	Pro	Thr	Pro	Val
				290					295					300
Pro	Arg	Leu	Ser	Glu	Ser	Ala	Asn	Pro	Pro	Ala	Gln	Arg	Pro	Asp
				305					310					315
Pro	Ala	His	Pro	Gly	Gly	Pro	Lys	Pro	Gln	Lys	Thr	Glu	Glu	Glu
				320					325					330
Val	Leu	Glu	Thr	Glu	Gly	Asp	Gln	Glu	Val	Ser	Leu	Gly	Thr	Pro
				335					340					345
Gln	Glu	Val	Val	Glu	Ala	Pro	Glu	Thr	Pro	Gly	Glu	Pro	Pro	Leu
				350					355					360

Ser Pro Gly Phe

&lt;210&gt; 19

&lt;211&gt; 152

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7503139CD1

&lt;400&gt; 19

Met	Glu	Arg	Leu	Lys	Ser	His	Lys	Pro	Ala	Thr	Met	Thr	Ile	Met
1				5					10					15
Val	Glu	Asp	Ile	Met	Lys	Leu	Leu	Cys	Ser	Leu	Ser	Gly	Glu	Arg
				20					25					30
Lys	Met	Lys	Ala	Ala	Val	Lys	His	Ser	Gly	Lys	Gly	Ala	Leu	Val
				35					40					45
Thr	Gly	Ala	Met	Ala	Phe	Val	Gly	Gly	Leu	Val	Gly	Gly	Pro	Pro
				50					55					60
Gly	Leu	Ala	Val	Gly	Gly	Ala	Val	Gly	Gly	Leu	Leu	Gly	Ala	Trp
				65					70					75
Met	Thr	Ser	Gly	Gln	Phe	Lys	Pro	Val	Pro	Gln	Ile	Leu	Met	Glu
				80					85					90
Leu	Pro	Pro	Ala	Glu	Gln	Gln	Arg	Leu	Phe	Asn	Glu	Ala	Ala	Ala
				95					100					105
Ile	Ile	Arg	His	Leu	Glu	Trp	Thr	Asp	Ala	Val	Gln	Leu	Thr	Ala
				110					115					120
Leu	Val	Met	Gly	Ser	Glu	Ala	Leu	Gln	Gln	Gln	Leu	Leu	Ala	Met
				125					130					135
Leu	Val	Asn	Tyr	Val	Thr	Lys	Glu	Leu	Arg	Ala	Glu	Ile	Gln	Tyr
				140					145					150

Asp Asp

&lt;210&gt; 20

&lt;211&gt; 756

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7505836CD1

&lt;400&gt; 20

Met	Ser	Phe	Leu	Gly	Ile	Leu	Cys	Lys	Cys	Pro	Leu	Gln	Asn	Glu
1				5					10					15
Ser	Gln	Glu	Glu	Ser	Tyr	Asn	Ala	Tyr	Pro	Leu	Pro	Ala	Val	Lys
				20					25					30
Val	Ser	Met	Asp	Trp	Leu	Arg	Leu	Arg	Pro	Arg	Val	Phe	Gln	Glu
				35					40					45
Ala	Val	Val	Asp	Glu	Arg	Gln	Tyr	Ile	Trp	Pro	Trp	Leu	Ile	Ser
				50					55					60
Leu	Leu	Asn	Ser	Phe	His	Pro	His	Glu	Glu	Asp	Leu	Ser	Ser	Ile
				65					70					75
Ser	Ala	Thr	Pro	Leu	Pro	Glu	Glu	Phe	Glu	Leu	Gln	Gly	Phe	Leu
				80					85					90
Ala	Leu	Arg	Pro	Ser	Phe	Arg	Asn	Leu	Asp	Phe	Ser	Lys	Gly	His
				95					100					105
Gln	Gly	Ile	Thr	Gly	Asp	Lys	Glu	Gly	Gln	Gln	Arg	Arg	Ile	Arg
				110					115					120
Gln	Gln	Arg	Leu	Ile	Ser	Ile	Gly	Lys	Trp	Ile	Ala	Asp	Asn	Gln
				125					130					135
Pro	Arg	Leu	Ile	Gln	Cys	Glu	Asn	Glu	Val	Gly	Lys	Leu	Leu	Phe
				140					145					150
Ile	Thr	Glu	Ile	Pro	Glu	Leu	Ile	Leu	Glu	Asp	Pro	Ser	Glu	Ala
				155					160					165
Lys	Glu	Asn	Leu	Ile	Leu	Gln	Glu	Thr	Ser	Val	Ile	Glu	Ser	Leu
				170					175					180
Ala	Ala	Asp	Gly	Ser	Pro	Gly	Leu	Lys	Ser	Val	Leu	Ser	Thr	Ser
				185					190					195
Arg	Asn	Leu	Ser	Asn	Asn	Cys	Asp	Thr	Gly	Glu	Lys	Pro	Val	Val

	200		205		210
Thr Phe Lys Glu	Asn Ile Lys Thr Arg	Glu Val Asn Arg Asp	Gln		
	215		220		225
Gly Arg Ser Phe	Pro Pro Lys Glu Val	Lys Ser Gln Thr Glu	Leu		
	230		235		240
Arg Lys Thr Pro	Val Ser Glu Ala Arg	Lys Thr Pro Val Thr	Gln		
	245		250		255
Thr Pro Thr Gln	Ala Ser Asn Ser Gln	Phe Ile Pro Ile His	His		
	260		265		270
Pro Gly Ala Phe	Pro Pro Leu Pro Ser	Arg Pro Gly Phe Pro	Pro		
	275		280		285
Pro Thr Tyr Val	Ile Pro Pro Pro Val	Ala Phe Ser Met Gly	Ser		
	290		295		300
Gly Tyr Thr Phe	Pro Ala Gly Val Ser	Val Pro Gly Thr Phe	Leu		
	305		310		315
Gln Pro Thr Ala	His Ser Pro Ala Gly	Asn Gln Val Gln Ala	Gly		
	320		325		330
Lys Gln Ser His	Ile Pro Tyr Ser Gln	Gln Arg Pro Ser Gly	Pro		
	335		340		345
Gly Pro Met Asn	Gln Gly Pro Gln Gln	Ser Gln Pro Pro Ser	Gln		
	350		355		360
Gln Pro Leu Thr	Ser Leu Pro Ala Gln	Pro Thr Ala Gln Ser	Thr		
	365		370		375
Ser Gln Leu Gln	Val Gln Ala Leu Thr	Gln Gln Gln Gln Ser	Pro		
	380		385		390
Thr Lys Ala Val	Pro Ala Leu Gly Lys	Ser Pro Pro His His	Ser		
	395		400		405
Gly Phe Gln Gln	Tyr Gln Gln Ala Asp	Ala Ser Lys Gln Leu	Trp		
	410		415		420
Asn Pro Pro Gln	Val Gln Gly Pro Leu	Gly Lys Ile Met Pro	Val		
	425		430		435
Lys Gln Pro Tyr	Tyr Leu Gln Thr Gln	Asp Pro Ile Lys Leu	Phe		
	440		445		450
Glu Pro Ser Leu	Gln Pro Pro Val Met	Gln Gln Gln Pro Leu	Lys		
	455		460		465
Lys Lys Met Lys	Pro Phe Pro Met Glu	Pro Tyr Asn His Asn	Pro		
	470		475		480
Ser Glu Val Lys	Val Pro Glu Phe Tyr	Trp Asp Ser Ser Tyr	Ser		
	485		490		495
Met Ala Asp Asn	Arg Ser Val Met Ala	Gln Gln Ala Asn Ile	Asp		
	500		505		510
Arg Arg Gly Lys	Arg Ser Pro Gly Ile	Phe Arg Pro Glu Gln	Asp		
	515		520		525
Pro Val Pro Arg	Met Pro Phe Glu Lys	Ser Leu Leu Glu Lys	Pro		
	530		535		540
Ser Glu Leu Met	Ser His Ser Ser Ser	Phe Leu Ser Leu Thr	Gly		
	545		550		555
Phe Ser Leu Asn	Gln Glu Arg Tyr Pro	Asn Asn Ser Met Phe	Asn		
	560		565		570
Glu Val Tyr Gly	Lys Asn Leu Thr Ser	Ser Ser Lys Ala Glu	Leu		
	575		580		585
Ser Pro Ser Met	Ala Pro Gln Glu Thr	Ser Leu Tyr Ser Leu	Phe		
	590		595		600
Glu Gly Thr Pro	Trp Ser Pro Ser Leu	Pro Ala Ser Ser Asp	His		
	605		610		615
Ser Thr Pro Ala	Ser Gln Ser Pro His	Ser Ser Asn Pro Ser	Ser		
	620		625		630
Leu Pro Ser Ser	Pro Pro Thr His Asn	His Asn Ser Val Pro	Phe		
	635		640		645
Ser Asn Phe Gly	Pro Ile Gly Thr Pro	Asp Asn Arg Asp Arg	Arg		
	650		655		660
Thr Ala Asp Arg	Trp Lys Thr Asp Lys	Pro Ala Met Gly Gly	Phe		
	665		670		675

Gly	Ile	Asp	Tyr	Leu	Ser	Ala	Thr	Ser	Ser	Ser	Glu	Ser	Ser	Trp
				680					685					690
His	Gln	Ala	Ser	Thr	Pro	Ser	Gly	Thr	Trp	Thr	Gly	His	Gly	Pro
				695					700					705
Ser	Met	Glu	Asp	Ser	Ser	Ala	Val	Leu	Met	Glu	Ser	Leu	Lys	Ser
				710					715					720
Ile	Trp	Ser	Ser	Ser	Met	Met	His	Pro	Gly	Pro	Ser	Ala	Leu	Glu
				725					730					735
Gln	Leu	Leu	Met	Gln	Gln	Lys	Gln	Lys	Gln	Gln	Arg	Gly	Gln	Gly
				740					745					750
Thr	Met	Asn	Pro	Pro	His									
				755										

<210> 21  
 <211> 120  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7505858CD1

<400> 21														
Met	Ala	Arg	His	Val	Phe	Leu	Thr	Gly	Pro	Pro	Gly	Val	Gly	Lys
1				5					10					15
Thr	Thr	Leu	Ile	His	Lys	Ala	Ser	Glu	Val	Leu	Lys	Ser	Ser	Gly
				20					25					30
Val	Pro	Val	Asp	Gly	Phe	Tyr	Thr	Glu	Glu	Val	Arg	Gln	Gly	Gly
				35					40					45
Arg	Arg	Ile	Gly	Phe	Asp	Val	Val	Thr	Leu	Ser	Gly	Thr	Arg	Gly
				50					55					60
Pro	Leu	Ser	Arg	Val	Gly	Leu	Glu	Pro	Pro	Pro	Gly	Lys	Arg	Glu
				65					70					75
Cys	Arg	Val	Gly	Gln	Tyr	Val	Val	Asp	Leu	Thr	Ser	Phe	Glu	Gln
				80					85					90
Leu	Ala	Leu	Pro	Val	Leu	Arg	Asn	Val	Thr	Lys	Glu	Asn	Arg	Asn
				95					100					105
His	Leu	Leu	Pro	Asp	Ile	Val	Thr	Cys	Val	Gln	Ser	Ser	Arg	Lys
				110					115					120

<210> 22  
 <211> 328  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7505872CD1

<400> 22														
Met	Pro	Gly	Ile	Ser	Ala	Arg	Gly	Leu	Ser	His	Glu	Gly	Arg	Lys
1				5					10					15
Gln	Leu	Ala	Val	Asn	Leu	Thr	Arg	Val	Leu	Ala	Leu	Tyr	Arg	Ser
				20					25					30
Ile	Leu	Asp	Ala	Tyr	Ile	Ile	Glu	Phe	Phe	Thr	Asp	Asn	Leu	Trp
				35					40					45
Asp	Thr	Leu	Pro	Cys	Ser	Trp	Gln	Glu	Ala	Leu	Asp	Gly	Leu	Lys
				50					55					60
Pro	Pro	Gln	Leu	Ala	Thr	Met	Leu	Leu	Gly	Met	Pro	Gly	Glu	Gly
				65					70					75
Glu	Val	Val	Arg	Tyr	Arg	Ser	Val	Trp	Pro	Leu	Thr	Leu	Leu	Ala
				80					85					90

```

Leu Lys Ser Thr Ala Cys Ala Leu Ala Phe Thr Arg Met Pro Gly
      95      100      105
Phe Gln Thr Pro Ser Glu Phe Leu Glu Asn Pro Ser Gln Ser Ser
      110      115      120
Arg Leu Thr Ala Pro Phe Arg Lys His Val Arg Pro Lys Lys Gln
      125      130      135
His Glu Ile Arg Arg Leu Gly Glu Leu Val Lys Lys Leu Ser Asp
      140      145      150
Phe Thr Gly Cys Thr Gln Val Val Asp Val Gly Ser Gly Gln Gly
      155      160      165
His Leu Ser Arg Phe Met Ala Leu Gly Leu Gly Leu Met Val Lys
      170      175      180
Ser Ile Glu Gly Asp Gln Arg Leu Val Glu Arg Ala Gln Arg Leu
      185      190      195
Asp Gln Glu Leu Leu Gln Ala Leu Glu Lys Glu Glu Lys Arg Asn
      200      205      210
Pro Gln Val Val Gln Thr Ser Pro Arg His Ser Pro His His Val
      215      220      225
Val Arg Tyr Val Gln Arg Gly Leu Gln Arg Val Gly Leu Asp Pro
      230      235      240
Gln Leu Pro Leu Asn Leu Ala Ala Leu Gln Ala His Leu Ala Gln
      245      250      255
Glu Asn Arg Val Val Ala Phe Phe Ser Leu Ala Leu Leu Leu Ala
      260      265      270
Pro Leu Val Glu Thr Leu Ile Leu Leu Asp Arg Leu Leu Tyr Leu
      275      280      285
Gln Glu Gln Gly Phe His Ala Glu Leu Leu Pro Ile Phe Ser Pro
      290      295      300
Glu Leu Ser Pro Arg Asn Leu Val Leu Val Ala Thr Lys Met Pro
      305      310      315
Leu Gly Gln Ala Leu Ser Val Leu Glu Thr Glu Asp Ser
      320      325

```

&lt;210&gt; 23

&lt;211&gt; 214

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7506456CD1

&lt;400&gt; 23

```

Met Ser Arg Ile Tyr His Asp Gly Ala Leu Arg Asn Lys Ala Val
  1      5      10      15
Gln Ser Val Arg Leu Pro Gly Ala Trp Asp Pro Ala Ala His Gln
      20      25      30
Gly Gly Asn Gly Val Leu Leu Glu Gly Glu Leu Ile Asp Val Ser
      35      40      45
Arg His Ser Ile Leu Asp Thr His Gly Arg Lys Glu Arg Tyr Tyr
      50      55      60
Val Leu Tyr Ile Arg Pro Ser His Ile His Arg Arg Lys Phe Asp
      65      70      75
Ala Lys Gly Asn Glu Ile Glu Pro Asn Phe Ser Ala Thr Arg Lys
      80      85      90
Val Asn Thr Gly Phe Leu Met Ser Ser Tyr Lys Val Glu Ala Lys
      95      100      105
Gly Leu Val Asn Lys Pro Glu Leu Leu Ala Leu Thr Glu Ser Leu
      110      115      120
Thr Pro Asp His Thr Val Ala Phe Trp Met Pro Glu Ser Glu Met
      125      130      135
Glu Val Met Glu Leu Glu Leu Gly Ala Gly Val Arg Leu Lys Thr
      140      145      150

```

Arg Gly Asp Gly	Pro Phe Leu Asp Ser	Leu Ala Lys Leu Glu Ala
	155	160
Gly Thr Val Thr	Lys Cys Asn Phe Thr	Gly Asp Gly Lys Thr Gly
	170	175
Ala Ser Trp Thr	Asp Asn Ile Met Ala	Gln Lys Cys Ser Lys Gly
	185	190
Ala Ala Ala Glu	Ile Arg Glu Gln Gly	Asp Gly Ala Glu Asp Glu
	200	205
Glu Trp Asp Asp		210

<210> 24  
 <211> 1442  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7506697CD1

<400> 24

Met Ser Ile Ala Ile	Pro Leu Gly Val Thr	Thr Ser Asp Thr Ser
1	5	10
Tyr Ser Asp Met Ala	Ala Gly Ser Asp Pro	Glu Ser Val Glu Ala
	20	25
Ser Pro Ala Val Asn	Glu Lys Ser Val Tyr	Ser Thr His Asn Tyr
	35	40
Gly Thr Thr Gln Arg	His Gly Cys Arg Gly	Leu Pro Tyr Ala Thr
	50	55
Ile Ile Pro Arg Ser	Asp Leu Asn Gly Leu	Pro Ser Pro Val Glu
	65	70
Glu Arg Cys Gly Asp	Ser Pro Asn Ser Glu	Gly Glu Thr Val Pro
	80	85
Thr Trp Cys Pro Cys	Gly Leu Ser Gln Asp	Gly Phe Leu Leu Asn
	95	100
Cys Asp Lys Cys Arg	Gly Met Ser Arg Gly	Lys Val Ile Arg Leu
	110	115
His Arg Arg Lys Gln	Asp Asn Ile Ser Gly	Gly Asp Ser Ser Ala
	125	130
Thr Glu Ser Trp Asp	Glu Glu Leu Ser Pro	Ser Thr Val Leu Tyr
	140	145
Thr Ala Thr Gln His	Thr Pro Thr Ser Ile	Thr Leu Thr Val Arg
	155	160
Arg Thr Lys Pro Lys	Lys Arg Lys Lys Ser	Pro Glu Lys Gly Arg
	170	175
Ala Ala Pro Lys Thr	Lys Lys Ile Lys Asn	Ser Pro Ser Glu Ala
	185	190
Gln Asn Leu Asp Glu	Asn Thr Thr Glu Trp	Glu Asn Arg Ile
	200	205
Arg Leu Trp Thr Asp	Gln Tyr Glu Glu Ala	Phe Thr Asn Gln Tyr
	215	220
Ser Ala Asp Val Gln	Asn Ala Leu Glu Gln	His Leu His Ser Ser
	230	235
Lys Glu Phe Val Gly	Lys Pro Thr Ile Leu	Asp Thr Ile Asn Lys
	245	250
Thr Glu Leu Ala Cys	Asn Asn Thr Val Ile	Gly Ser Gln Met Gln
	260	265
Leu Gln Leu Gly Arg	Val Thr Arg Val Gln	Lys His Arg Lys Ile
	275	280
Leu Arg Ala Ala Arg	Asp Leu Ala Leu Asp	Thr Leu Ile Ile Glu
	290	295
Tyr Arg Gly Lys Val	Met Leu Arg Gln Gln	Phe Glu Val Asn Gly
	305	310

His	Phe	Phe	Lys	Lys	Pro	Tyr	Pro	Phe	Val	Leu	Phe	Tyr	Ser	Lys
				320					325					330
Phe	Asn	Gly	Val	Glu	Met	Cys	Val	Asp	Ala	Arg	Thr	Phe	Gly	Asn
				335					340					345
Asp	Ala	Arg	Phe	Ile	Arg	Arg	Ser	Cys	Thr	Pro	Asn	Ala	Glu	Val
				350					355					360
Arg	His	Met	Ile	Ala	Asp	Gly	Met	Ile	His	Leu	Cys	Ile	Tyr	Ala
				365					370					375
Val	Ser	Ala	Ile	Thr	Lys	Asp	Ala	Glu	Val	Thr	Ile	Ala	Phe	Asp
				380					385					390
Tyr	Glu	Tyr	Ser	Asn	Cys	Asn	Tyr	Lys	Val	Asp	Cys	Ala	Cys	His
				395					400					405
Lys	Gly	Asn	Arg	Asn	Cys	Pro	Ile	Gln	Lys	Arg	Asn	Pro	Asn	Ala
				410					415					420
Thr	Glu	Leu	Pro	Leu	Leu	Pro	Pro	Pro	Pro	Ser	Leu	Pro	Thr	Ile
				425					430					435
Gly	Ala	Glu	Thr	Arg	Arg	Arg	Lys	Ala	Arg	Arg	Lys	Glu	Leu	Glu
				440					445					450
Met	Glu	Gln	Gln	Asn	Glu	Ala	Ser	Glu	Glu	Asn	Asn	Asp	Gln	Gln
				455					460					465
Ser	Gln	Glu	Val	Pro	Glu	Lys	Val	Thr	Val	Ser	Ser	Asp	His	Glu
				470					475					480
Glu	Val	Asp	Asn	Pro	Glu	Glu	Lys	Pro	Glu	Glu	Glu	Lys	Glu	Glu
				485					490					495
Val	Ile	Asp	Asp	Gln	Glu	Asn	Leu	Ala	His	Ser	Arg	Arg	Thr	Arg
				500					505					510
Glu	Asp	Arg	Lys	Val	Glu	Ala	Ile	Met	His	Ala	Phe	Glu	Asn	Leu
				515					520					525
Glu	Lys	Arg	Lys	Lys	Arg	Arg	Asp	Gln	Pro	Leu	Glu	Gln	Ser	Asn
				530					535					540
Ser	Asp	Val	Glu	Ile	Thr	Thr	Thr	Thr	Ser	Glu	Thr	Pro	Val	Gly
				545					550					555
Glu	Glu	Thr	Lys	Thr	Glu	Ala	Pro	Glu	Ser	Glu	Val	Ser	Asn	Ser
				560					565					570
Val	Ser	Asn	Val	Thr	Ile	Pro	Ser	Thr	Pro	Gln	Ser	Val	Gly	Val
				575					580					585
Asn	Thr	Arg	Arg	Ser	Ser	Gln	Ala	Gly	Asp	Ile	Ala	Ala	Glu	Lys
				590					595					600
Leu	Val	Pro	Lys	Pro	Pro	Pro	Ala	Lys	Pro	Ser	Arg	Pro	Arg	Pro
				605					610					615
Lys	Ser	Arg	Ile	Ser	Arg	Tyr	Arg	Thr	Ser	Ser	Ala	Gln	Arg	Leu
				620					625					630
Lys	Arg	Gln	Lys	Gln	Ala	Asn	Ala	Gln	Gln	Ala	Glu	Leu	Ser	Gln
				635					640					645
Ala	Ala	Leu	Glu	Glu	Gly	Gly	Ser	Asn	Ser	Leu	Val	Thr	Pro	Thr
				650					655					660
Glu	Ala	Gly	Ser	Leu	Asp	Ser	Ser	Gly	Glu	Asn	Arg	Pro	Leu	Thr
				665					670					675
Gly	Ser	Asp	Pro	Thr	Val	Val	Ser	Ile	Thr	Gly	Ser	His	Val	Asn
				680					685					690
Arg	Ala	Ala	Ser	Lys	Tyr	Pro	Lys	Thr	Lys	Lys	Tyr	Leu	Val	Thr
				695					700					705
Glu	Trp	Leu	Asn	Asp	Lys	Ala	Glu	Lys	Gln	Glu	Cys	Pro	Val	Glu
				710					715					720
Cys	Pro	Leu	Arg	Ile	Thr	Thr	Asp	Pro	Thr	Val	Leu	Ala	Thr	Thr
				725					730					735
Leu	Asn	Met	Leu	Pro	Gly	Leu	Ile	His	Ser	Pro	Leu	Ile	Cys	Thr
				740					745					750
Thr	Pro	Lys	His	Tyr	Ile	Arg	Phe	Gly	Ser	Pro	Phe	Ile	Pro	Glu
				755					760					765
Arg	Arg	Arg	Arg	Pro	Leu	Leu	Pro	Asp	Gly	Thr	Phe	Ser	Ser	Cys
				770					775					780
Lys	Lys	Arg	Trp	Ile	Lys	Gln	Ala	Leu	Glu	Glu	Gly	Met	Thr	Gln



	785		790		795
Thr Ser Ser Val	Pro Gln Glu Thr Arg	Thr Gln His Leu Tyr	Gln		
	800		805		810
Ser Asn Glu Asn	Ser Ser Ser Ser Ser	Ile Cys Lys Asp Asn	Ala		
	815		820		825
Asp Leu Leu Ser	Pro Leu Lys Lys Trp	Lys Ser Arg Tyr Leu	Met		
	830		835		840
Glu Gln Asn Val	Thr Lys Leu Leu Arg	Pro Leu Ser Pro Val	Thr		
	845		850		855
Pro Pro Pro Pro	Asn Ser Gly Ser Lys	Ser Pro Gln Leu Ala	Thr		
	860		865		870
Pro Gly Ser Ser	His Pro Gly Glu Glu	Glu Cys Arg Asn Gly	Tyr		
	875		880		885
Ser Leu Met Phe	Ser Pro Val Thr Ser	Leu Thr Thr Ala Ser	Arg		
	890		895		900
Cys Asn Thr Pro	Leu Gln Phe Glu Leu	Cys His Arg Lys Asp	Leu		
	905		910		915
Asp Leu Ala Lys	Val Gly Tyr Leu Asp	Ser Asn Thr Asn Ser	Cys		
	920		925		930
Ala Asp Arg Pro	Ser Leu Leu Asn Ser	Gly His Ser Asp Leu	Ala		
	935		940		945
Pro His Pro Ser	Leu Gly Pro Thr Ser	Glu Thr Gly Phe Pro	Ser		
	950		955		960
Arg Ser Gly Asp	Gly His Gln Thr Leu	Val Arg Asn Ser Asp	Gln		
	965		970		975
Ala Phe Arg Thr	Glu Phe Asn Leu Met	Tyr Ala Tyr Ser Pro	Leu		
	980		985		990
Asn Ala Met Pro	Arg Ala Asp Gly Leu	Tyr Arg Gly Ser Pro	Leu		
	995		1000		1005
Val Gly Asp Arg	Lys Pro Leu His Leu	Asp Gly Gly Tyr Cys	Ser		
	1010		1015		1020
Pro Ala Glu Gly	Phe Ser Ser Arg Tyr	Glu His Gly Leu Met	Lys		
	1025		1030		1035
Asp Leu Ser Arg	Gly Ser Leu Ser Pro	Gly Gly Glu Arg Ala	Cys		
	1040		1045		1050
Glu Gly Val Pro	Ser Ala Pro Gln Asn	Pro Pro Gln Arg Lys	Lys		
	1055		1060		1065
Val Ser Leu Leu	Glu Tyr Arg Lys Arg	Lys Gln Glu Ala Lys	Glu		
	1070		1075		1080
Asn Ser Ala Gly	Gly Gly Gly Asp Ser	Ala Gln Ser Lys Ser	Lys		
	1085		1090		1095
Ser Ala Gly Ala	Gly Gln Gly Ser Ser	Asn Ser Val Ser Asp	Thr		
	1100		1105		1110
Gly Ala His Gly	Val Gln Gly Ser Ser	Ala Arg Thr Pro Ser	Ser		
	1115		1120		1125
Pro His Lys Lys	Phe Ser Pro Ser His	Ser Ser Met Ser His	Leu		
	1130		1135		1140
Glu Ala Val Ser	Pro Ser Asp Ser Arg	Gly Thr Ser Ser Ser	His		
	1145		1150		1155
Cys Arg Pro Gln	Glu Asn Ile Ser Ser	Arg Trp Met Val Pro	Thr		
	1160		1165		1170
Ser Val Glu Arg	Leu Arg Glu Gly Gly	Ser Ile Pro Lys Val	Leu		
	1175		1180		1185
Arg Ser Ser Val	Arg Val Ala Gln Lys	Gly Glu Pro Ser Pro	Thr		
	1190		1195		1200
Trp Glu Ser Asn	Ile Thr Glu Lys Asp	Ser Asp Pro Ala Asp	Gly		
	1205		1210		1215
Glu Gly Pro Glu	Thr Leu Ser Ser Ala	Leu Ser Lys Gly Ala	Thr		
	1220		1225		1230
Val Tyr Ser Pro	Ser Arg Tyr Ser Tyr	Gln Leu Leu Gln Cys	Asp		
	1235		1240		1245
Ser Pro Arg Thr	Glu Ser Gln Ser Leu	Leu Gln Gln Ser Ser	Ser		
	1250		1255		1260

```

Pro Phe Arg Gly His Pro Thr Gln Ser Pro Gly Tyr Ser Tyr Arg
      1265      1270      1275
Thr Thr Ala Leu Arg Pro Gly Asn Pro Pro Ser His Gly Ser Ser
      1280      1285      1290
Glu Ser Ser Leu Ser Ser Thr Ser Tyr Ser Ser Pro Ala His Pro
      1295      1300      1305
Val Ser Thr Asp Ser Leu Ala Pro Phe Thr Gly Thr Pro Gly Tyr
      1310      1315      1320
Phe Ser Ser Gln Pro His Ser Gly Asn Ser Thr Gly Ser Asn Leu
      1325      1330      1335
Pro Arg Arg Ser Cys Pro Ser Ser Ala Ala Ser Pro Thr Leu Gln
      1340      1345      1350
Gly Pro Ser Asp Ser Pro Thr Ser Asp Ser Val Ser Gln Ser Ser
      1355      1360      1365
Thr Gly Thr Leu Ser Ser Thr Ser Phe Pro Gln Asn Ser Arg Ser
      1370      1375      1380
Ser Leu Pro Ser Asp Leu Arg Thr Ile Ser Leu Pro Ser Ala Gly
      1385      1390      1395
Gln Ser Ala Val Tyr Gln Ala Ser Arg Val Ser Ala Val Ser Asn
      1400      1405      1410
Ser Gln His Tyr Pro His Arg Gly Ser Gly Gly Val His Gln Tyr
      1415      1420      1425
Arg Leu Gln Pro Leu Gln Gly Ser Gly Val Lys Thr Gln Thr Gly
      1430      1435      1440
Leu Ser

```

<210> 25

<211> 268

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7623472CD1

<400> 25

```

Met Ser Gly Lys Ala Asn Ala Ser Lys Lys Asn Ala Gln Gln Leu
  1      5      10      15
Lys Arg Asn Pro Lys Arg Lys Lys Asp Asn Glu Glu Val Val Leu
      20      25      30
Ser Glu Asn Lys Val Arg Asn Thr Val Lys Lys Asn Lys Asn His
      35      40      45
Leu Lys Asp Leu Ser Ser Glu Gly Gln Thr Lys His Thr Asn Leu
      50      55      60
Lys His Gly Lys Thr Ala Ala Ser Lys Arg Lys Thr Trp Gln Pro
      65      70      75
Leu Ser Lys Ser Thr Arg Asp His Leu Gln Thr Met Met Glu Ser
      80      85      90
Val Ile Met Thr Ile Leu Ser Asn Ser Ile Lys Glu Lys Glu Glu
      95      100      105
Ile Gln Tyr His Leu Asn Phe Leu Lys Lys Arg Leu Leu Gln Gln
      110      115      120
Cys Glu Thr Leu Lys Val Pro Pro Lys Lys Met Glu Asp Leu Thr
      125      130      135
Asn Val Ser Ser Leu Leu Asn Met Glu Arg Ala Arg Asp Lys Ala
      140      145      150
Asn Glu Glu Gly Leu Ala Leu Leu Gln Glu Glu Ile Asp Lys Met
      155      160      165
Val Glu Thr Thr Glu Leu Met Thr Gly Asn Ile Gln Ser Leu Lys
      170      175      180
Asn Lys Ile Gln Ile Leu Ala Ser Glu Val Glu Glu Glu Glu Glu
      185      190      195

```

Arg	Val	Lys	Gln	Met	His	Gln	Ile	Asn	Ser	Ser	Gly	Val	Leu	Ser
				200					205					210
Leu	Pro	Glu	Leu	Ser	Gln	Lys	Thr	Leu	Lys	Ala	Pro	Thr	Leu	Gln
				215					220					225
Lys	Glu	Ile	Leu	Ala	Leu	Ile	Pro	Asn	Gln	Asn	Ala	Leu	Leu	Lys
				230					235					240
Asp	Leu	Asp	Ile	Leu	His	Asn	Ser	Ser	Gln	Met	Lys	Ser	Met	Ser
				245					250					255
Thr	Phe	Ile	Glu	Glu	Ala	Tyr	Lys	Lys	Leu	Asp	Ala	Ser		
				260					265					

&lt;210&gt; 26

&lt;211&gt; 588

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7506416CD1

&lt;400&gt; 26

Met	Leu	Ser	Leu	Cys	Ala	Gln	Asp	Leu	Thr	Gln	Met	Leu	Ala	Leu
1				5					10					15
Ser	Arg	His	Ser	Leu	Leu	Ser	Pro	Leu	Leu	Ser	Val	Thr	Ser	Phe
				20					25					30
Arg	Arg	Phe	Tyr	Arg	Gly	Asp	Ser	Pro	Thr	Asp	Ser	Gln	Lys	Asp
				35					40					45
Met	Ile	Glu	Ile	Pro	Leu	Pro	Pro	Trp	Gln	Glu	Arg	Thr	Asp	Glu
				50					55					60
Ser	Ile	Glu	Thr	Lys	Arg	Ala	Arg	Leu	Leu	Tyr	Glu	Ser	Arg	Lys
				65					70					75
Arg	Gly	Met	Leu	Glu	Asn	Cys	Ile	Leu	Leu	Ser	Leu	Phe	Ala	Lys
				80					85					90
Glu	His	Leu	Gln	His	Met	Thr	Glu	Lys	Gln	Leu	Asn	Leu	Tyr	Asp
				95					100					105
Arg	Leu	Ile	Asn	Glu	Pro	Ser	Asn	Asp	Trp	Asp	Ile	Tyr	Tyr	Trp
				110					115					120
Ala	Thr	Gly	Arg	Arg	Phe	Tyr	Thr	Arg	Lys	Trp	His	Ile	Leu	Lys
				125					130					135
Trp	Ser	Ser	Thr	Asp	Ser	Asn	Ser	Ser	Gln	Pro	Cys	Gly	Gly	Gly
				140					145					150
Arg	Arg	Leu	Gly	Pro	Glu	Pro	Trp	Lys	Gln	Gly	Leu	Ala	Arg	Ala
				155					160					165
Ala	Ser	Asp	Pro	Pro	Leu	Leu	Ala	Arg	Pro	Pro	Gly	Ala	Leu	Pro
				170					175					180
His	Ser	Ile	Met	Met	Gly	Lys	Leu	Pro	Leu	Gly	Val	Val	Ser	Pro
				185					190					195
Tyr	Val	Lys	Met	Ser	Ser	Gly	Gly	Tyr	Thr	Asp	Pro	Leu	Lys	Phe
				200					205					210
Tyr	Ala	Thr	Ser	Tyr	Cys	Thr	Ala	Tyr	Gly	Arg	Glu	Asp	Phe	Lys
				215					220					225
Pro	Arg	Val	Gly	Ser	His	Val	Gly	Thr	Gly	Tyr	Lys	Ser	Asn	Phe
				230					235					240
Gln	Pro	Val	Val	Ser	Cys	Gln	Ala	Ser	Leu	Glu	Ala	Leu	Asp	Asn
				245					250					255
Pro	Ala	Arg	Gly	Glu	Gln	Ala	Gln	Asp	His	Phe	Gln	Ser	Val	Ala
				260					265					270
Ser	Gln	Ser	Tyr	Arg	Pro	Leu	Glu	Val	Pro	Asp	Gly	Lys	His	Pro
				275					280					285
Leu	Pro	Trp	Ser	Met	Arg	Gln	Thr	Ser	Ser	Gly	Tyr	Gly	Arg	Glu
				290					295					300
Lys	Pro	Ser	Ala	Gly	Pro	Pro	Thr	Lys	Glu	Val	Arg	Lys	Val	His
				305					310					315

Phe	Asp	Thr	Gln	Glu	His	Gly	Pro	Gln	Ala	Ile	Thr	Gly	Leu	Glu
				320					325					330
Pro	Arg	Glu	Val	Pro	Leu	Leu	His	Gln	Gln	Gln	Gly	Gln	Asp	Pro
				335					340					345
Leu	Glu	Arg	Glu	Asn	Phe	Arg	His	Gly	Pro	Arg	Phe	Met	Thr	Ser
				350					355					360
Glu	Tyr	Asn	Ser	Lys	Tyr	Leu	Arg	Asp	Pro	Leu	Asp	Gln	Pro	Asp
				365					370					375
Phe	Leu	Gln	Lys	Lys	Ser	Ile	Gly	Ala	Lys	Glu	Gly	Ser	Gly	Phe
				380					385					390
Thr	Lys	Gln	Ser	His	Gln	Ser	Pro	Ile	Val	Phe	Gln	Pro	Pro	Ser
				395					400					405
Gln	Ala	Leu	Pro	Gly	Asp	Pro	Gly	Asp	Glu	Phe	Leu	Pro	Val	Leu
				410					415					420
Ala	Arg	Gly	Ser	Lys	Arg	Glu	Thr	Ala	Phe	Ser	Arg	Gly	Asn	Glu
				425					430					435
Arg	Ile	Leu	Asn	Pro	Arg	Val	Pro	Pro	Pro	Cys	Pro	Glu	Pro	Ser
				440					445					450
Ser	Val	Ser	His	Gln	Gln	Phe	Gln	Pro	Leu	His	Arg	Met	Gln	Gln
				455					460					465
Thr	Asn	Val	Ala	Leu	Leu	Gly	Arg	Glu	Thr	Val	Gly	Lys	Lys	Glu
				470					475					480
Pro	Thr	Gly	Phe	Ser	Leu	Asn	Asn	Pro	Met	Tyr	Val	Arg	Ser	Pro
				485					490					495
Cys	Asp	Pro	Asp	Arg	Asp	Gln	Arg	Tyr	Leu	Thr	Thr	Tyr	Asn	Gln
				500					505					510
Gly	Tyr	Phe	Glu	Asn	Ile	Pro	Lys	Gly	Leu	Asp	Gln	Glu	Gly	Trp
				515					520					525
Thr	Arg	Gly	Gly	Ile	Gln	Pro	Gln	Met	Pro	Gly	Gly	Tyr	Ala	Leu
				530					535					540
Ser	Gln	Pro	Val	Ser	Cys	Met	Glu	Ala	Thr	Pro	Asn	Pro	Met	Glu
				545					550					555
Ser	Leu	Arg	His	Leu	His	Pro	His	Val	Gly	Arg	Thr	Leu	Thr	Ser
				560					565					570
Ala	Asp	Pro	Phe	Tyr	Gln	Asn	Thr	Pro	His	Ser	Ser	Arg	Cys	Val
				575					580					585
Ala	His	Ser												

&lt;210&gt; 27

&lt;211&gt; 306

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 4823849CD1

&lt;400&gt; 27

Met	Thr	Leu	Lys	Cys	Ile	Trp	Gly	Gly	Gln	Arg	Lys	Gly	Ser	Ile
1				5					10					15
Ala	Gly	Gln	Gly	Pro	Ala	Gln	Cys	Val	His	Thr	Cys	Pro	Gln	Asp
				20					25					30
Asp	Val	Pro	Leu	Ser	Ser	Ile	Cys	Ile	Pro	Pro	Leu	Val	Phe	Cys
				35					40					45
Ser	Phe	Leu	Thr	His	Val	Pro	Glu	Ala	Asp	Phe	Gln	Val	Thr	Lys
				50					55					60
Pro	Gly	Asn	Trp	Arg	Asp	Val	Cys	Glu	Gly	Ser	Ala	Thr	Val	Ile
				65					70					75
Leu	Gly	Val	Thr	Ser	Ser	Val	Pro	Ser	Leu	Pro	Leu	Pro	Asn	Val
				80					85					90
Leu	Leu	Met	Ala	Asn	Val	Thr	Trp	Pro	Gln	Gly	Pro	Phe	Thr	Thr
				95					100					105

```

Trp Ser Thr Pro Gly Asp Ala Pro Val Ile Asn Leu Ser Arg Leu
110 115 120
Leu Pro Leu Lys Tyr Val Glu Leu Arg Ile Tyr Asp Arg Leu Gln
125 130 135
Arg Ile Leu Arg Val Arg Thr Val Thr Glu Lys Ile Tyr Tyr Leu
140 145 150
Lys Leu His Glu Lys His Pro Glu Ile Val Phe Gln Phe Trp Val
155 160 165
Arg Leu Val Lys Ile Leu Gln Lys Gly Leu Ser Ile Thr Ile Lys
170 175 180
Asp Pro Arg Ile Lys Phe Thr His Cys Leu Val Pro Lys Met Pro
185 190 195
Thr Asn Ser Thr Glu Thr Thr Pro Glu Asn Ser Leu Leu Ser Ser
200 205 210
Pro Gln Pro Ser Glu Pro Leu Val Leu Leu Ala Ala Glu Gln Thr
215 220 225
Ser Gly Ser Phe Ser Gln Leu Ser Gly Lys Pro Gln Leu Thr Ala
230 235 240
Asp Arg Asn Asn Asp Thr Ala Ile Glu Ile Asp Asn Cys Ser Ser
245 250 255
Tyr Lys Ile Pro Ser Pro Val Ala Ser Pro Ile Asn Leu Asn Ile
260 265 270
Pro Met Arg Ala Ala Leu Ser His Ser Leu Trp Glu Gln Glu Asp
275 280 285
Trp Asn Glu His Leu Leu Gln Val His Ile Ala Ser Tyr Leu Gly
290 295 300
Glu His Phe Leu Gly Ala
305

```

```

<210> 28
<211> 239
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 4433922CD1

```

```

<400> 28
Met Arg Gln Pro Ser Pro Ile Val Pro Ala Leu Gln Asn Lys Ile
1 5 10 15
Ala Ser Lys Leu Gln Arg Pro Pro Ser Val Asp Ser Ile Ile Arg
20 25 30
Ser Phe Ile His Glu Ser Ser Met Ser Arg Ala Gln Ser Pro Pro
35 40 45
Val Pro Ala Arg Lys Asn Gln Leu Arg Ala Glu Glu Glu Lys Lys
50 55 60
Asn Val Ile Met Glu Leu Ser Glu Met Arg Lys Gln Leu Arg Ser
65 70 75
Glu Glu Arg Arg Leu Gln Glu Arg Leu Leu His Met Asp Ser Asp
80 85 90
Asp Glu Ile Pro Ile Arg Lys Lys Glu Arg Asn Pro Met Asp Ile
95 100 105
Phe Asp Met Ala Arg His Arg Leu Gln Ala Pro Val Arg Arg Gln
110 115 120
Ser Pro Lys Gly Leu Asp Ala Ala Thr Phe Gln Asn Val His Asp
125 130 135
Phe Asn Glu Leu Lys Asp Arg Asp Ser Glu Thr Arg Val Asp Leu
140 145 150
Lys Phe Met Tyr Leu Asp Pro Pro Arg Asp His His Thr Leu Glu
155 160 165
Ile Gln Gln Gln Ala Leu Leu Arg Glu Gln Gln Lys Arg Leu Asn
170 175 180

```

Arg	Ile	Lys	Met	Gln	Glu	Gly	Ala	Lys	Val	Asp	Leu	Asp	Ala	Ile
				185					190					195
Pro	Ser	Ala	Lys	Val	Arg	Glu	Gln	Arg	Met	Gly	Leu	Thr	Val	Arg
				200					205					210
His	Ile	Leu	Pro	Leu	Lys	Met	Thr	Ser	Ser	Leu	His	His	His	Ser
				215					220					225
Cys	Pro	Leu	His	Gly	Ser	Ala	Gly	Gly	Thr	Asn	Gly	Lys	Asp	
				230					235					

&lt;210&gt; 29

&lt;211&gt; 49

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7504597CD1

&lt;400&gt; 29

Met	Phe	Pro	Glu	Pro	Pro	Thr	Pro	Gly	Pro	Pro	Ser	Pro	Asp	Thr
1				5					10					15
Pro	Pro	Asp	Ser	Ser	Arg	Ile	Ser	His	Gly	Pro	Gly	Ala	Ala	Arg
				20					25					30
Ser	Arg	Gly	Ala	Gly	Ala	Ser	Thr	Ile	Arg	Ala	Arg	Ala	Ala	Gly
				35					40					45
Gly	Arg	Gln	Ala											

&lt;210&gt; 30

&lt;211&gt; 247

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7505987CD1

&lt;400&gt; 30

Met	Ala	Ser	Ala	Glu	Leu	Asp	Tyr	Thr	Ile	Glu	Ile	Pro	Asp	Gln
1				5					10					15
Pro	Cys	Trp	Ser	Gln	Lys	Asn	Ser	Pro	Ser	Pro	Gly	Gly	Lys	Glu
				20					25					30
Ala	Glu	Thr	Arg	Gln	Pro	Val	Val	Ile	Leu	Leu	Gly	Trp	Gly	Gly
				35					40					45
Cys	Lys	Asp	Lys	Asn	Leu	Ala	Lys	Tyr	Ser	Ala	Ile	Tyr	His	Lys
				50					55					60
Arg	Lys	Leu	Leu	Glu	Leu	Leu	Phe	Asp	Tyr	Glu	Ile	Glu	Lys	Glu
				65					70					75
Pro	Leu	Leu	Phe	His	Val	Phe	Ser	Asn	Gly	Gly	Val	Met	Leu	Tyr
				80					85					90
Arg	Tyr	Val	Leu	Glu	Leu	Leu	Gln	Thr	Arg	Arg	Phe	Cys	Arg	Leu
				95					100					105
Arg	Val	Val	Gly	Thr	Ile	Phe	Asp	Ser	Ala	Pro	Gly	Asp	Ser	Asn
				110					115					120
Leu	Val	Gly	Ala	Leu	Arg	Ala	Leu	Ala	Ala	Ile	Leu	Glu	Arg	Arg
				125					130					135
Ala	Ala	Met	Leu	Arg	Leu	Leu	Leu	Leu	Val	Ala	Phe	Ala	Leu	Val
				140					145					150
Val	Val	Leu	Phe	His	Val	Leu	Leu	Ala	Pro	Ile	Thr	Ala	Leu	Phe
				155					160					165
His	Thr	His	Phe	Tyr	Asp	Arg	Leu	Gln	Asp	Ala	Gly	Ser	Arg	Trp
				170					175					180
Pro	Glu	Leu	Tyr	Leu	Tyr	Ser	Arg	Ala	Asp	Glu	Val	Val	Leu	Ala

	185		190		195
Arg Asp Ile Glu	Arg Met Val Glu Ala	Arg Leu Ala Arg Arg	Val		
	200		205		210
Leu Ala Arg Ser	Val Asp Phe Val Ser	Ser Ala His Val Ser	His		
	215		220		225
Leu Arg Asp Tyr	Pro Thr Tyr Tyr Thr	Ser Leu Cys Val Asp	Phe		
	230		235		240
Met Arg Asn Cys	Val Arg Cys				
	245				

&lt;210&gt; 31

&lt;211&gt; 418

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7506025CD1

&lt;400&gt; 31

Met Ser Glu Leu Thr	Lys Glu Leu Met Glu	Leu Val Trp Gly Thr
1	5	10
Lys Ser Ser Pro Gly	Leu Ser Asp Thr Ile	Phe Cys Arg Trp Thr
	20	25
Gln Gly Phe Val Phe	Ser Glu Ser Glu Gly	Ser Ala Leu Glu Gln
	35	40
Phe Glu Gly Gly Pro	Cys Ala Val Ile Ala	Pro Val Gln Ala Phe
	50	55
Leu Leu Lys Lys Leu	Leu Phe Ser Ser Glu	Lys Ser Ser Trp Arg
	65	70
Asp Cys Ser Glu Glu	Gln Lys Glu Leu	Cys His Thr Leu
	80	85
Cys Asp Ile Leu Glu	Ser Ala Cys Cys Asp	His Ser Gly Ser Tyr
	95	100
Cys Leu Val Ser Trp	Leu Arg Gly Lys Thr	Thr Glu Glu Thr Ala
	110	115
Ser Ile Ser Gly Ser	Pro Ala Glu Ser Ser	Cys Gln Val Glu His
	125	130
Ser Ser Ala Leu Ala	Val Glu Glu Leu Gly	Phe Glu Arg Phe His
	140	145
Ala Leu Ile Gln Lys	Arg Ser Phe Arg Ser	Leu Pro Glu Leu Lys
	155	160
Asp Ala Val Leu Asp	Gln Tyr Ser Met Trp	Gly Asn Lys Phe Gly
	170	175
Val Leu Leu Phe Leu	Tyr Ser Val Leu Leu	Thr Lys Gly Ile Glu
	185	190
Asn Ile Lys Asn Glu	Ile Glu Asp Ala Ser	Glu Pro Leu Ile Asp
	200	205
Pro Val Tyr Gly His	Gly Ser Gln Ser Leu	Ile Asn Leu Leu Leu
	215	220
Thr Gly His Ala Val	Ser Asn Val Trp Asp	Gly Asp Arg Glu Cys
	230	235
Ser Gly Met Lys Leu	Leu Gly Ile His Glu	Gln Ala Ala Val Gly
	245	250
Phe Leu Thr Leu Met	Glu Ala Leu Arg Tyr	Cys Lys Asp Met Ala
	260	265
Leu Val Ala Pro Glu	Ala Pro Ser Glu Gln	Ala Arg Arg Val Phe
	275	280
Gln Thr Tyr Asp Pro	Glu Asp Asn Gly Phe	Ile Pro Asp Ser Leu
	290	295
Leu Glu Asp Val Met	Lys Ala Leu Asp Leu	Val Ser Asp Pro Glu
	305	310
Tyr Ile Asn Leu Met	Lys Asn Lys Leu Asp	Pro Glu Gly Leu Gly

	320		325		330
Ile Ile Leu Leu Gly	Pro Phe Leu Gln	Glu Phe Phe Pro Asp	Gln		
	335		340		345
Gly Ser Ser Gly	Pro Glu Ser Phe Thr	Val Tyr His Tyr Asn	Gly		
	350		355		360
Leu Lys Gln Ser Asn	Tyr Asn Glu Lys	Val Met Tyr Val Glu	Gly		
	365		370		375
Thr Ala Val Val Met	Gly Phe Glu Asp	Pro Met Leu Gln Thr	Asp		
	380		385		390
Asp Thr Pro Ile Lys	Arg Cys Leu Gln	Thr Lys Trp Pro Tyr	Ile		
	395		400		405
Glu Leu Leu Trp Thr	Thr Asp Arg Ser	Pro Ser Leu Asn			
	410		415		

<210> 32  
 <211> 139  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7506102CD1

<400> 32	
Met Cys Phe Pro Lys	Val Leu Ser Asp Asp Met Lys Lys Leu Lys
1	5 10 15
Ala Arg Met Glu Leu	Thr Pro Leu Leu Glu Lys Glu Arg Asp Gly
	20 25 30
Leu Arg Cys Arg Gly	Asn Arg Ser Pro Val Pro Asp Val Glu Asp
	35 40 45
Pro Ala Thr Glu Glu	Pro Gly Glu Ser Phe Cys Asp Lys Val Met
	50 55 60
Arg Trp Phe Gln Ala	Met Leu Gln Arg Leu Gln Thr Trp Trp His
	65 70 75
Gly Val Leu Ala Trp	Val Lys Glu Lys Val Val Ala Leu Val His
	80 85 90
Ala Val Gln Ala Leu	Trp Lys Gln Phe Gln Ser Phe Cys Cys Ser
	95 100 105
Leu Ser Glu Leu Phe	Met Ser Ser Phe Gln Ser Tyr Gly Ala Pro
	110 115 120
Arg Gly Asp Lys Glu	Glu Leu Thr Pro Gln Lys Cys Ser Glu Pro
	125 130 135
Gln Ser Ser Lys	

<210> 33  
 <211> 295  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1333949CD1

<400> 33	
Met Ser Ser Asn Thr	Met Leu Gln Lys Thr Leu Leu Ile Leu Ile
1	5 10 15
Ser Phe Ser Val Val	Thr Trp Met Ile Phe Ile Ile Ser Gln Asn
	20 25 30
Phe Thr Lys Leu Trp	Ser Ala Leu Asn Leu Ser Ile Ser Val His
	35 40 45
Tyr Trp Asn Asn Ser	Ala Lys Ser Leu Phe Pro Lys Thr Ser Leu
	50 55 60



```

Ile Pro Leu Lys Pro Leu Thr Glu Thr Glu Leu Arg Ile Lys Glu
      65      70
Ile Ile Glu Lys Leu Asp Gln Gln Ile Pro Pro Arg Pro Phe Thr
      80      85
His Val Asn Thr Thr Thr Ser Ala Thr His Ser Thr Ala Thr Ile
      95     100
Leu Asn Pro Arg Asp Thr Tyr Cys Arg Gly Asp Gln Leu Asp Ile
     110     115
Leu Leu Glu Val Arg Asp His Leu Gly Gln Arg Lys Gln Tyr Gly
     125     130
Gly Asp Phe Leu Arg Ala Arg Met Ser Ser Pro Ala Leu Thr Ala
     140     145
Gly Ala Ser Gly Lys Val Met Asp Phe Asn Asn Gly Thr Tyr Leu
     155     160
Val Ser Phe Thr Leu Phe Trp Glu Gly Gln Val Ser Leu Ser Leu
     170     175
Leu Leu Ile His Pro Ser Glu Gly Ala Ser Ala Leu Trp Arg Ala
     185     190
Arg Asn Gln Gly Tyr Asp Lys Ile Ile Phe Lys Gly Lys Phe Val
     200     205
Asn Gly Thr Ser His Val Phe Thr Glu Cys Gly Leu Thr Leu Asn
     215     220
Ser Asn Ala Glu Leu Cys Glu Tyr Leu Asp Asp Arg Asp Gln Glu
     230     235
Ala Phe Tyr Cys Met Lys Pro Gln His Met Pro Cys Glu Ala Leu
     245     250
Thr Tyr Met Thr Thr Arg Asn Arg Glu Val Ser Tyr Leu Thr Asp
     260     265
Lys Glu Asn Ser Leu Phe His Arg Met Ala Pro Gly Glu Thr Ser
     275     280
Ile Ala Gly Asn Gln Val Gln Ser Gly Ser
     290     295

```

&lt;210&gt; 34

&lt;211&gt; 540

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7035533CD1

&lt;400&gt; 34

```

Met Ala Val Glu Cys Ala Glu Ile Arg Arg His His Arg Val Gly
  1      5      10
Ile Lys Asp Ile Ala Gly Ile His Leu Pro Thr Asn Val Lys Phe
     20     25
Gln Ser Pro Ala Tyr Ser Ser Val Asp Thr Glu Glu Thr Ile Glu
     35     40
Pro Tyr Thr Thr Glu Lys Met Ser Arg Val Pro Gly Gly Tyr Leu
     50     55
Ala Leu Thr Glu Cys Phe Glu Ile Met Thr Val Asp Phe Asn Asn
     65     70
Leu Gln Glu Leu Lys Ser Leu Ala Thr Lys Lys Pro Asp Lys Ile
     80     85
Gly Ile Pro Val Ile Lys Glu Gly Ile Leu Asp Ala Ile Met Val
     95    100
Trp Phe Val Leu Gln Leu Asp Asp Glu His Ser Leu Ser Thr Ser
    110    115
Pro Ser Glu Glu Thr Cys Trp Glu Gln Ala Val Tyr Pro Val Gln
    125    130
Asp Leu Ala Asp Tyr Trp Ile Lys Pro Gly Asp His Val Met Met
    140    145

```

Glu	Val	Ser	Cys	Gln	Asp	Cys	Tyr	Leu	Arg	Ile	Gln	Ser	Ile	Ser	155	160	165
Val	Leu	Gly	Leu	Glu	Cys	Glu	Met	Asp	Val	Ala	Lys	Ser	Phe	Thr	170	175	180
Gln	Asn	Lys	Asp	Leu	Leu	Ser	Leu	Gly	Asn	Glu	Ala	Glu	Leu	Cys	185	190	195
Ser	Ala	Leu	Ala	Asn	Leu	Gln	Thr	Ser	Lys	Pro	Asp	Ala	Val	Glu	200	205	210
Gln	Thr	Cys	Ile	Leu	Glu	Ser	Thr	Glu	Ile	Ala	Leu	Leu	Asn	Asn	215	220	225
Ile	Pro	Tyr	His	Glu	Gly	Phe	Lys	Met	Ala	Met	Ser	Lys	Val	Leu	230	235	240
Ser	Ser	Leu	Thr	Pro	Glu	Lys	Leu	Tyr	Gln	Thr	Met	Asp	Thr	His	245	250	255
Cys	Gln	Asn	Glu	Met	Ser	Ser	Gly	Thr	Gly	Gln	Ser	Asn	Thr	Val	260	265	270
Gln	Asn	Ile	Leu	Glu	Pro	Phe	Tyr	Val	Leu	Asp	Val	Ser	Glu	Gly	275	280	285
Phe	Ser	Val	Leu	Pro	Val	Ile	Ala	Gly	Thr	Leu	Gly	Gln	Val	Lys	290	295	300
Pro	Tyr	Ser	Ser	Val	Glu	Lys	Asp	Gln	His	Arg	Ile	Ala	Leu	Asp	305	310	315
Leu	Ile	Ser	Glu	Ala	Asn	His	Phe	Pro	Lys	Glu	Thr	Leu	Glu	Phe	320	325	330
Trp	Leu	Arg	His	Val	Glu	Asp	Glu	Ser	Ala	Met	Leu	Gln	Arg	Pro	335	340	345
Lys	Ser	Asp	Lys	Leu	Trp	Ser	Ile	Ile	Ile	Leu	Asp	Val	Ile	Glu	350	355	360
Pro	Ser	Gly	Leu	Ile	Gln	Gln	Glu	Ile	Met	Glu	Lys	Ala	Ala	Ile	365	370	375
Ser	Arg	Cys	Leu	Leu	Gln	Ser	Gly	Gly	Lys	Ile	Phe	Pro	Gln	Tyr	380	385	390
Val	Leu	Met	Phe	Gly	Leu	Leu	Val	Glu	Ser	Gln	Thr	Leu	Leu	Glu	395	400	405
Glu	Asn	Ala	Val	Gln	Gly	Thr	Glu	Arg	Thr	Leu	Gly	Leu	Asn	Ile	410	415	420
Ala	Pro	Phe	Ile	Asn	Gln	Phe	Gln	Val	Pro	Ile	Arg	Val	Phe	Leu	425	430	435
Asp	Leu	Ser	Ser	Leu	Pro	Cys	Ile	Pro	Leu	Ser	Lys	Pro	Val	Glu	440	445	450
Leu	Leu	Arg	Leu	Asp	Leu	Met	Thr	Pro	Tyr	Leu	Asn	Thr	Ser	Asn	455	460	465
Arg	Glu	Val	Lys	Val	Tyr	Val	Cys	Lys	Ser	Gly	Arg	Leu	Thr	Ala	470	475	480
Ile	Pro	Phe	Trp	Tyr	His	Met	Tyr	Leu	Asp	Glu	Glu	Ile	Arg	Leu	485	490	495
Asp	Thr	Ser	Ser	Glu	Ala	Ser	His	Trp	Lys	Gln	Ala	Ala	Val	Val	500	505	510
Leu	Asp	Asn	Pro	Ile	Gln	Val	Glu	Met	Gly	Glu	Glu	Leu	Val	Leu	515	520	525
Ser	Ile	Gln	His	His	Lys	Ser	Asn	Val	Ser	Ile	Thr	Val	Lys	Gln	530	535	540

&lt;210&gt; 35

&lt;211&gt; 791

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2815375CD1

&lt;400&gt; 35

Met	Ala	Gly	Thr	Ser	Ala	Pro	Gly	Ser	Lys	Arg	Arg	Ser	Glu	Pro
1				5					10					15
Pro	Ala	Pro	Arg	Pro	Gly	Pro	Pro	Pro	Gly	Thr	Gly	His	Pro	Pro
				20					25					30
Ser	Lys	Arg	Ala	Arg	Gly	Phe	Ser	Ala	Ala	Ala	Ala	Pro	Asp	Pro
				35					40					45
Asp	Asp	Pro	Phe	Gly	Ala	His	Gly	Asp	Phe	Thr	Ala	Asp	Asp	Leu
				50					55					60
Glu	Glu	Leu	Asp	Thr	Leu	Ala	Ser	Gln	Ala	Leu	Ser	Gln	Cys	Pro
				65					70					75
Ala	Ala	Ala	Arg	Asp	Val	Ser	Ser	Asp	His	Lys	Val	His	Arg	Leu
				80					85					90
Leu	Asp	Gly	Met	Ser	Lys	Asn	Pro	Ser	Gly	Lys	Asn	Arg	Glu	Thr
				95					100					105
Val	Pro	Ile	Lys	Asp	Asn	Phe	Glu	Leu	Glu	Val	Leu	Gln	Ala	Gln
				110					115					120
Tyr	Lys	Glu	Leu	Lys	Glu	Lys	Met	Lys	Val	Met	Glu	Glu	Glu	Val
				125					130					135
Leu	Ile	Lys	Asn	Gly	Glu	Ile	Lys	Ile	Leu	Arg	Asp	Ser	Leu	His
				140					145					150
Gln	Thr	Glu	Ser	Val	Leu	Glu	Glu	Gln	Arg	Arg	Ser	His	Phe	Leu
				155					160					165
Leu	Glu	Gln	Glu	Lys	Thr	Gln	Ala	Leu	Ser	Asp	Lys	Glu	Lys	Glu
				170					175					180
Phe	Ser	Lys	Lys	Leu	Gln	Ser	Leu	Gln	Ser	Glu	Leu	Gln	Phe	Lys
				185					190					195
Asp	Ala	Glu	Met	Asn	Glu	Leu	Arg	Thr	Lys	Leu	Gln	Thr	Ser	Glu
				200					205					210
Arg	Ala	Asn	Lys	Leu	Ala	Ala	Pro	Ser	Val	Ser	His	Val	Ser	Pro
				215					220					225
Arg	Lys	Asn	Pro	Ser	Val	Val	Ile	Lys	Pro	Glu	Ala	Cys	Ser	Pro
				230					235					240
Gln	Phe	Gly	Lys	Thr	Ser	Phe	Pro	Thr	Lys	Glu	Ser	Phe	Ser	Ala
				245					250					255
Asn	Met	Ser	Leu	Pro	His	Pro	Cys	Gln	Thr	Glu	Ser	Gly	Tyr	Lys
				260					265					270
Pro	Leu	Val	Gly	Arg	Glu	Asp	Ser	Lys	Pro	His	Ser	Leu	Arg	Gly
				275					280					285
Asp	Ser	Ile	Lys	Gln	Glu	Glu	Ala	Gln	Lys	Ser	Phe	Val	Asp	Ser
				290					295					300
Trp	Arg	Gln	Arg	Ser	Asn	Thr	Gln	Gly	Ser	Ile	Leu	Ile	Asn	Leu
				305					310					315
Leu	Leu	Lys	Gln	Pro	Leu	Ile	Pro	Gly	Ser	Ser	Leu	Ser	Leu	Cys
				320					325					330
His	Leu	Leu	Ser	Ser	Ser	Ser	Glu	Ser	Pro	Ala	Gly	Thr	Pro	Leu
				335					340					345
Gln	Pro	Pro	Gly	Phe	Gly	Ser	Thr	Leu	Ala	Gly	Met	Ser	Gly	Leu
				350					355					360
Arg	Thr	Thr	Gly	Ser	Tyr	Asp	Gly	Ser	Phe	Ser	Leu	Ser	Ala	Leu
				365					370					375
Arg	Glu	Ala	Gln	Asn	Leu	Ala	Phe	Thr	Gly	Leu	Asn	Leu	Val	Ala
				380					385					390
Arg	Asn	Glu	Cys	Ser	Arg	Asp	Gly	Asp	Pro	Ala	Glu	Gly	Gly	Arg
				395					400					405
Arg	Ala	Phe	Pro	Leu	Cys	Gln	Leu	Ser	Gly	Ala	Val	His	Phe	Leu
				410					415					420
Pro	Leu	Val	Gln	Phe	Phe	Ile	Gly	Leu	His	Cys	Gln	Ala	Leu	Gln
				425					430					435
Asp	Leu	Ala	Ala	Ala	Lys	Arg	Ser	Gly	Ala	Pro	Gly	Asp	Ser	Pro
				440					445					450
Thr	His	Ser	Ser	Cys	Val	Ser	Ser	Gly	Val	Glu	Thr	Asn	Pro	Glu
				455					460					465

```

Asp Ser Val Cys Ile Leu Glu Gly Phe Ser Val Thr Ala Leu Ser
470 475 480
Ile Leu Gln His Leu Val Cys His Ser Gly Ala Val Val Ser Leu
485 490 495
Leu Leu Ser Gly Val Gly Ala Asp Ser Ala Ala Gly Glu Gly Asn
500 505 510
Arg Ser Leu Val His Arg Leu Ser Asp Gly Asp Met Thr Ser Ala
515 520 525
Leu Arg Gly Val Ala Asp Asp Gln Gly Gln His Pro Leu Leu Lys
530 535 540
Met Leu Leu His Leu Leu Ala Phe Ser Ser Ala Ala Thr Gly His
545 550 555
Leu Gln Ala Ser Val Leu Thr Gln Cys Leu Lys Val Leu Val Lys
560 565 570
Leu Ala Glu Asn Thr Ser Cys Asp Phe Leu Pro Arg Phe Gln Cys
575 580 585
Val Phe Gln Val Leu Pro Lys Cys Leu Ser Pro Glu Thr Pro Leu
590 595 600
Pro Ser Val Leu Leu Ala Val Glu Leu Leu Ser Leu Leu Ala Asp
605 610 615
His Asp Gln Leu Ala Pro Gln Leu Cys Ser His Ser Glu Gly Cys
620 625 630
Leu Leu Leu Leu Leu Tyr Met Tyr Ile Thr Ser Arg Pro Asp Arg
635 640 645
Val Ala Leu Glu Thr Gln Trp Leu Gln Leu Glu Gln Glu Val Val
650 655 660
Trp Leu Leu Ala Lys Leu Gly Val Gln Ser Pro Leu Pro Pro Val
665 670 675
Thr Gly Ser Asn Cys Gln Cys Asn Val Glu Val Val Arg Ala Leu
680 685 690
Thr Val Met Leu His Arg Gln Trp Leu Thr Val Arg Arg Ala Gly
695 700 705
Gly Pro Pro Arg Thr Asp Gln Gln Arg Arg Thr Val Arg Cys Leu
710 715 720
Arg Asp Thr Val Leu Leu Leu His Gly Leu Ser Gln Lys Asp Lys
725 730 735
Leu Phe Met Met His Cys Val Glu Val Leu His Gln Phe Asp Gln
740 745 750
Val Met Pro Gly Val Ser Met Leu Ile Arg Gly Leu Pro Asp Val
755 760 765
Thr Asp Cys Glu Glu Ala Ala Leu Asp Asp Leu Cys Ala Ala Glu
770 775 780
Thr Asp Val Glu Asp Pro Glu Val Glu Cys Gly
785 790

```

&lt;210&gt; 36

&lt;211&gt; 154

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2820152CD1

&lt;400&gt; 36

```

Met Ser Ala Thr Arg Ala Lys Lys Val Lys Met Ala Thr Lys Ser
1 5 10 15
Cys Pro Glu Cys Asp Gln Gln Val Pro Val Ala Cys Lys Ser Cys
20 25 30
Pro Cys Gly Tyr Ile Phe Ile Ser Arg Lys Leu Leu Asn Ala Lys
35 40 45
His Ser Glu Lys Ser Pro Pro Ser Thr Glu Asn Lys His Glu Ala
50 55 60

```

```

Lys Arg Arg Arg Thr Glu Arg Val Arg Arg Glu Lys Ile Asn Ser
      65      70      75
Thr Val Asn Lys Asp Leu Glu Asn Arg Lys Arg Ser Arg Ser Asn
      80      85      90
Ser His Ser Asp His Ile Arg Arg Gly Arg Gly Arg Pro Lys Ser
      95     100     105
Ala Ser Ala Lys Lys His Glu Glu Glu Arg Glu Lys Gln Glu Lys
     110     115     120
Glu Ile Asp Ile Tyr Ala Asn Leu Ser Asp Glu Lys Ala Phe Val
     125     130     135
Phe Ser Val Ala Leu Ala Glu Ile Asn Arg Lys Ile Ile Asn Gln
     140     145     150
Arg Leu Ile Leu

```

```

<210> 37
<211> 957
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2959305CD1

```

```

<400> 37
Met Asp Asp Ala Ala Glu Cys Phe Glu Asn Met Leu Glu Arg Ile
 1      5      10      15
His Phe His Ile Val Pro Ser Arg Asp Ala Asp Met Cys Thr Ser
      20      25      30
Lys Ser Cys Ile Thr His Gln Lys Phe Ala Met Thr Leu Tyr Glu
      35      40      45
Gln Cys Val Cys Arg Ser Cys Gly Ala Ser Ser Asp Pro Leu Pro
      50      55      60
Phe Thr Glu Phe Val Arg Tyr Ile Ser Thr Thr Ala Leu Cys Asn
      65      70      75
Glu Val Glu Arg Met Leu Glu Arg His Glu Arg Phe Lys Pro Glu
      80      85      90
Met Phe Ala Glu Leu Leu Gln Ala Ala Asn Thr Thr Asp Asp Tyr
      95     100     105
Arg Lys Cys Pro Ser Asn Cys Gly Gln Lys Ile Lys Ile Arg Arg
     110     115     120
Val Leu Met Asn Cys Pro Glu Ile Val Thr Ile Gly Leu Val Trp
     125     130     135
Asp Ser Glu His Ser Asp Leu Thr Glu Ala Val Val Arg Asn Leu
     140     145     150
Ala Thr His Leu Tyr Leu Pro Gly Leu Phe Tyr Arg Val Thr Asp
     155     160     165
Glu Asn Ala Lys Asn Ser Glu Leu Asn Leu Val Gly Met Ile Cys
     170     175     180
Tyr Thr Ser Gln His Tyr Cys Ala Phe Ala Phe His Thr Lys Ser
     185     190     195
Ser Lys Trp Val Phe Phe Asp Asp Ala Asn Val Lys Glu Ile Gly
     200     205     210
Thr Arg Trp Lys Asp Val Val Ser Lys Cys Ile Arg Cys His Phe
     215     220     225
Gln Pro Leu Leu Leu Phe Tyr Ala Asn Pro Asp Gly Thr Ala Val
     230     235     240
Ser Thr Glu Asp Ala Leu Arg Gln Val Ile Ser Trp Ser His Tyr
     245     250     255
Lys Ser Val Ala Glu Asn Met Gly Cys Glu Lys Pro Val Ile His
     260     265     270
Lys Ser Asp Asn Leu Lys Glu Asn Gly Phe Gly Asp Gln Ala Lys
     275     280     285

```

Gln	Arg	Glu	Asn	Gln	Lys	Phe	Pro	Thr	Asp	Asn	Ile	Ser	Ser	Ser
				290					295					300
Asn	Arg	Ser	His	Ser	His	Thr	Gly	Val	Gly	Lys	Gly	Pro	Ala	Lys
				305					310					315
Leu	Ser	His	Ile	Asp	Gln	Arg	Glu	Lys	Ile	Lys	Asp	Ile	Ser	Arg
				320					325					330
Glu	Cys	Ala	Leu	Lys	Ala	Ile	Glu	Gln	Lys	Asn	Leu	Leu	Ser	Ser
				335					340					345
Gln	Arg	Lys	Asp	Leu	Glu	Lys	Gly	Gln	Arg	Lys	Asp	Leu	Gly	Arg
				350					355					360
His	Arg	Asp	Leu	Val	Asp	Glu	Asp	Leu	Ser	His	Phe	Gln	Ser	Gly
				365					370					375
Ser	Pro	Pro	Ala	Pro	Asn	Gly	Phe	Lys	Gln	His	Gly	Asn	Pro	His
				380					385					390
Leu	Tyr	His	Ser	Gln	Gly	Lys	Gly	Ser	Tyr	Lys	His	Asp	Arg	Val
				395					400					405
Val	Pro	Gln	Ser	Arg	Ala	Ser	Ala	Gln	Ile	Ile	Ser	Ser	Ser	Lys
				410					415					420
Ser	Gln	Ile	Leu	Ala	Pro	Gly	Glu	Lys	Ile	Thr	Gly	Lys	Val	Lys
				425					430					435
Ser	Asp	Asn	Gly	Thr	Gly	Tyr	Asp	Thr	Asp	Ser	Ser	Gln	Asp	Ser
				440					445					450
Arg	Asp	Arg	Gly	Asn	Ser	Cys	Asp	Ser	Ser	Ser	Lys	Ser	Arg	Asn
				455					460					465
Arg	Gly	Trp	Lys	Pro	Met	Arg	Glu	Thr	Leu	Asn	Val	Asp	Ser	Ile
				470					475					480
Phe	Ser	Glu	Ser	Glu	Lys	Arg	Gln	His	Ser	Pro	Arg	His	Lys	Pro
				485					490					495
Asn	Ile	Ser	Asn	Lys	Pro	Lys	Ser	Ser	Lys	Asp	Pro	Ser	Phe	Ser
				500					505					510
Asn	Trp	Pro	Lys	Glu	Asn	Pro	Lys	Gln	Lys	Gly	Leu	Met	Thr	Ile
				515					520					525
Tyr	Glu	Asp	Glu	Met	Lys	Gln	Glu	Ile	Gly	Ser	Arg	Ser	Ser	Leu
				530					535					540
Glu	Ser	Asn	Gly	Lys	Gly	Ala	Glu	Lys	Asn	Lys	Gly	Leu	Val	Glu
				545					550					555
Gly	Lys	Val	His	Gly	Asp	Asn	Trp	Gln	Met	Gln	Arg	Thr	Glu	Ser
				560					565					570
Gly	Tyr	Glu	Ser	Ser	Asp	His	Ile	Ser	Asn	Gly	Ser	Thr	Asn	Leu
				575					580					585
Asp	Ser	Pro	Val	Ile	Asp	Gly	Asn	Gly	Thr	Val	Met	Asp	Ile	Ser
				590					595					600
Gly	Val	Lys	Glu	Thr	Val	Cys	Phe	Ser	Asp	Gln	Ile	Thr	Thr	Ser
				605					610					615
Asn	Leu	Asn	Lys	Glu	Arg	Gly	Asp	Cys	Thr	Ser	Leu	Gln	Ser	Gln
				620					625					630
His	His	Leu	Glu	Gly	Phe	Arg	Lys	Glu	Leu	Arg	Asn	Leu	Glu	Ala
				635					640					645
Gly	Tyr	Lys	Ser	His	Glu	Phe	His	Pro	Glu	Ser	His	Leu	Gln	Ile
				650					655					660
Lys	Asn	His	Leu	Ile	Lys	Arg	Ser	His	Val	His	Glu	Asp	Asn	Gly
				665					670					675
Lys	Leu	Phe	Pro	Ser	Ser	Ser	Leu	Gln	Ile	Pro	Lys	Asp	His	Asn
				680					685					690
Ala	Arg	Glu	His	Ile	His	Gln	Ser	Asp	Glu	Gln	Lys	Leu	Glu	Lys
				695					700					705
Pro	Asn	Glu	Cys	Lys	Phe	Ser	Glu	Trp	Leu	Asn	Ile	Glu	Asn	Ser
				710					715					720
Glu	Arg	Thr	Gly	Leu	Pro	Phe	His	Val	Asp	Asn	Ser	Ala	Ser	Gly
				725					730					735
Lys	Arg	Val	Asn	Ser	Asn	Glu	Pro	Ser	Ser	Leu	Trp	Ser	Ser	His
				740					745					750
Leu	Arg	Thr	Val	Gly	Leu	Lys	Pro	Glu	Thr	Ala	Pro	Leu	Ile	Gln

	755		760		765
Gln Gln Asn Ile	Met Asp Gln Cys Tyr	Phe Glu Asn Ser Leu	Ser		
	770		775		780
Thr Glu Cys Ile	Ile Arg Ser Ala Ser	Arg Ser Asp Gly Cys	Gln		
	785		790		795
Met Pro Lys Leu	Phe Cys Gln Asn Leu	Pro Pro Pro Leu Pro	Pro		
	800		805		810
Lys Lys Tyr Ala	Ile Thr Ser Val Pro	Gln Ser Glu Lys Ser	Glu		
	815		820		825
Ser Thr Pro Asp	Val Lys Leu Thr Glu	Val Phe Lys Ala Thr	Ser		
	830		835		840
His Leu Pro Lys	His Ser Leu Ser Thr	Ala Ser Glu Pro Ser	Leu		
	845		850		855
Glu Val Ser Thr	His Met Asn Asp Glu	Arg His Lys Glu Thr	Phe		
	860		865		870
Gln Val Arg Glu	Cys Phe Gly Asn Thr	Pro Asn Cys Pro Ser	Ser		
	875		880		885
Ser Ser Thr Asn	Asp Phe Gln Ala Asn	Ser Gly Ala Ile Asp	Ala		
	890		895		900
Phe Cys Gln Pro	Glu Leu Asp Ser Ile	Ser Thr Cys Pro Asn	Glu		
	905		910		915
Thr Val Ser Leu	Thr Thr Tyr Phe Ser	Val Asp Ser Cys Met	Thr		
	920		925		930
Asp Thr Tyr Arg	Leu Lys Tyr His Gln	Arg Pro Lys Leu Ser	Phe		
	935		940		945
Pro Glu Ser Ser	Gly Phe Cys Asn Asn	Ser Leu Ser			
	950		955		

&lt;210&gt; 38

&lt;211&gt; 340

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 4913449CD1

&lt;400&gt; 38

Met Ser Ser Pro Lys	Ala Glu Leu Lys	Leu Glu Lys Lys	Ser Gly	
1	5	10	15	
Arg Lys Pro Arg Ser	Pro Arg Asp Ser	Gly Pro Gln Lys	Glu Leu	
	20	25	30	
Val Ile Pro Gly Ile	Val Asp Phe Glu	Arg Ile Arg Arg	Ala Leu	
	35	40	45	
Arg Thr Pro Lys Pro	Gln Thr Pro Gly	Thr Tyr Cys Phe	Gly Arg	
	50	55	60	
Leu Ser His His Ser	Phe Phe Ser Arg	His His Pro His	Pro Gln	
	65	70	75	
His Val Thr His Ile	Gln Asp Leu Thr	Gly Lys Pro Val	Cys Val	
	80	85	90	
Val Arg Asp Phe Pro	Ala Pro Leu Pro	Glu Ser Thr Val	Phe Ser	
	95	100	105	
Gly Cys Gln Met Gly	Ile Pro Thr Ile	Ser Val Pro Ile	Gly Asp	
	110	115	120	
Pro Gln Ser Asn Arg	Asn Pro Gln Leu	Ser Ser Glu Ala	Trp Lys	
	125	130	135	
Lys Glu Leu Lys Glu	Leu Ala Ser Arg	Val Ala Phe Leu	Thr Lys	
	140	145	150	
Glu Asp Glu Leu Lys	Lys Lys Glu Lys	Glu Gln Lys Glu	Glu Pro	
	155	160	165	
Leu Arg Glu Gln Gly	Ala Lys Tyr Ser	Ala Glu Thr Gly	Arg Leu	
	170	175	180	
Ile Pro Ala Ser Thr	Arg Ala Val Gly	Arg Arg Arg Ser	His Gln	

	185		190		195
Gly Gln Gln Ser	Gln Ser Ser Ser Arg	His Glu Gly Val Gln	Ala		
	200		205		210
Phe Leu Leu Gln	Asp Gln Glu Leu Leu	Val Leu Glu Leu Leu	Cys		
	215		220		225
Arg Ile Leu Glu	Thr Asp Leu Leu Ser	Ala Ile Gln Phe Trp	Leu		
	230		235		240
Leu Tyr Ala Pro	Pro Lys Glu Lys Asp	Leu Ala Leu Gly Leu	Leu		
	245		250		255
Gln Thr Ala Val	Ala Gln Leu Leu Pro	Gln Pro Leu Val Ser	Ile		
	260		265		270
Pro Thr Glu Lys	Leu Leu Ser Gln Leu	Pro Glu Val His Glu	Pro		
	275		280		285
Pro Gln Glu Lys	Gln Glu Pro Pro Cys	Ser Gln Ser Pro Lys	Lys		
	290		295		300
Thr Lys Ile Ser	Pro Phe Thr Lys Ser	Glu Lys Pro Glu Tyr	Ile		
	305		310		315
Gly Glu Ala Gln	Val Leu Gln Met His	Ser Ser Gln Asn Thr	Glu		
	320		325		330
Lys Lys Thr Ser	Lys Pro Arg Ala Glu	Ser			
	335		340		

&lt;210&gt; 39

&lt;211&gt; 287

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7506136CD1

&lt;400&gt; 39

Met Trp Trp Gly Gly	Arg Gly Gln Ser	Phe Asn Ile Ala Pro	Gln	
1	5	10	15	
Lys Glu Glu Pro Glu	Met Gly Ser Val	Gln Glu Asn Arg Met	Pro	
	20	25	30	
Glu Pro Arg Ser Arg	Gln Pro Ser Ser	Cys Leu Ala Ser Arg	Cys	
	35	40	45	
Leu Pro Gly Glu Gln	Ile Leu Ala Trp	Ala Pro Gly Val Arg	Lys	
	50	55	60	
Gly Leu Glu Pro Glu	Leu Ser Gly Thr	Leu Ile Cys Thr Asn	Phe	
	65	70	75	
Arg Val Thr Phe Gln	Pro Cys Gly Trp	Gln Trp Asn Gln Asp	Thr	
	80	85	90	
Pro Leu Asn Ser Glu	Tyr Asp Phe Ala	Leu Val Asn Ile Gly	Arg	
	95	100	105	
Leu Glu Ala Val Ser	Gly Leu Ser Arg	Val Gln Leu Leu Arg	Pro	
	110	115	120	
Gly Ser Leu His Lys	Phe Ile Pro Glu	Glu Ile Leu Ile His	Gly	
	125	130	135	
Arg Asp Phe Arg Leu	Leu Arg Val Gly	Phe Glu Ala Gly Gly	Leu	
	140	145	150	
Glu Pro Gln Ala Phe	Gln Val Thr Met	Ala Ile Val Gln Ala	Arg	
	155	160	165	
Ala Gln Ser Asn Gln	Ala Gln Gln Tyr	Ser Gly Ile Thr Leu	Ser	
	170	175	180	
Lys Ala Gly Gln Gly	Ser Gly Ser Arg	Lys Pro Pro Ile Pro	Leu	
	185	190	195	
Met Glu Thr Ala Glu	Asp Trp Glu Thr	Glu Arg Lys Lys Gln	Ala	
	200	205	210	
Ala Arg Gly Trp Arg	Val Ser Thr Val	Asn Glu Arg Phe Asp	Val	
	215	220	225	
Ala Thr Ser Leu Pro	Arg Tyr Phe Trp	Val Pro Asn Arg Ile	Leu	



	230		235		240
Asp Ser Glu Val	Arg Arg Ala Phe Gly	His Phe His Gln Gly	Arg		
	245		250		255
Gly Pro Arg Leu	Ser Trp His His Pro	Gly Gly Ser Asp Leu	Leu		
	260		265		270
Arg Cys Gly Gly	Phe Tyr Thr Ala Ser	Asp Pro Asn Lys Glu	Asp		
	275		280		285
Ile Arg					

<210> 40  
 <211> 139  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7506225CD1

<400> 40	
Met Cys Phe Pro Lys Val Leu Ser Asp Asp	Met Lys Lys Leu Lys
1 5	10 15
Ala Arg Met Glu Leu Thr Pro Leu Leu Glu	Lys Glu Arg Asp Gly
20 25	30
Leu Arg Cys Arg Gly Asn Arg Ser Pro Val	Pro Asp Val Glu Asp
35 40	45
Pro Ala Thr Glu Glu Pro Gly Glu Ser Phe	Cys Asp Lys Val Met
50 55	60
Arg Trp Phe Gln Ala Met Leu Gln Arg Leu	Gln Thr Trp Trp His
65 70	75
Gly Val Leu Ala Trp Val Lys Glu Lys Val	Val Ala Leu Val His
80 85	90
Ala Val Gln Ala Leu Trp Lys Gln Phe Gln	Ser Phe Cys Cys Ser
95 100	105
Leu Ser Glu Leu Phe Met Ser Ser Phe Gln	Ser Tyr Gly Ala Pro
110 115	120
Arg Gly Asp Lys Glu Leu Thr Pro Gln Lys	Cys Ser Glu Pro
125 130	135
Gln Ser Ser Lys	

<210> 41  
 <211> 185  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7506227CD1

<400> 41	
Met Ala Met Phe Glu Gln Met Arg Ala Asn	Val Gly Lys Leu Leu
1 5	10 15
Lys Gly Ile Asp Arg Tyr Gln Phe Asn Pro	Ala Phe Phe Gln Thr
20 25	30
Thr Val Thr Ala Gln Ile Leu Leu Lys Ala	Leu Thr Asn Leu Pro
35 40	45
His Thr Asp Phe Thr Leu Cys Lys Cys Met	Ile Asp Gln Ala His
50 55	60
Gln Glu Glu Arg Pro Ile Arg Gln Ile Leu	Tyr Leu Gly Asp Leu
65 70	75
Leu Glu Thr Cys His Phe Gln Ala Phe Trp	Gln Ala Leu Asp Glu
80 85	90

```

Asn Met Asp Leu Leu Glu Gly Ile Thr Gly Phe Glu Asp Ser Val
      95      100
Arg Lys Phe Ile Cys His Val Val Gly Ile Thr Tyr Gln His Ile
      110      115
Asp Arg Trp Leu Leu Ala Glu Met Leu Gly Asp Leu Ser Asp Ser
      125      130
Gln Leu Lys Val Trp Met Ser Lys Tyr Gly Trp Ser Ala Asp Glu
      140      145
Ser Gly Gln Ile Phe Ile Cys Ser Gln Glu Glu Ser Ile Lys Pro
      155      160
Lys Asn Ile Val Glu Lys Ile Asp Phe Asp Ser Val Ser Ser Ile
      170      175
Met Ala Ser Ser Gln
      185

```

<210> 42  
 <211> 164  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3144431CD1

```

<400> 42
Met Val Glu Ile Ala Gly Leu Asn Thr Asp Leu Gln Gln Glu Tyr
  1      5      10      15
Thr Arg Gln Arg Glu His Leu Glu Arg Asn Leu Ala Thr Leu Lys
      20      25      30
Lys Lys Val Val Lys Glu Gly Glu Leu His Arg Thr Asp Tyr Val
      35      40      45
Arg Ile Met Gln Glu Asn Val Ser Leu Ile Lys Glu Ile Asn Glu
      50      55      60
Leu Arg Arg Glu Leu Lys Phe Thr Arg Ser Gln Val Tyr Asp Leu
      65      70      75
Glu Ala Ala Leu Lys Leu Thr Lys Lys Val Arg Pro Gln Glu Val
      80      85      90
Ser Glu Thr Glu Pro Ser Arg Asp Met Leu Ser Thr Ala Pro Thr
      95     100     105
Ala Arg Leu Asn Glu Gln Glu Glu Thr Gly Arg Ile Ile Glu Met
     110     115     120
Gln Arg Leu Glu Ile Gln Arg Leu Arg Asp Gln Ile Gln Glu Gln
     125     130     135
Glu Gln Val Thr Gly Phe His Thr Leu Ala Gly Val Arg Leu Pro
     140     145     150
Ser Leu Ser Asn Ser Glu Val Asp Leu Glu Val Lys Thr Asn
     155     160

```

<210> 43  
 <211> 577  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2633315CD1

```

<400> 43
Met Leu Thr Pro Ala Phe Asp Leu Ser Gln Asp Pro Asp Phe Leu
  1      5      10      15
Thr Ile Ala Ile Arg Val Pro Tyr Ala Arg Val Ser Glu Phe Asp
      20      25      30
Val Tyr Phe Glu Gly Ser Asp Phe Lys Phe Tyr Ala Lys Pro Tyr

```

				35					40					45
Phe	Leu	Arg	Leu	Thr	Leu	Pro	Gly	Arg	Ile	Val	Glu	Asn	Gly	Ser
				50					55					60
Glu	Gln	Gly	Ser	Tyr	Asp	Ala	Asp	Lys	Gly	Ile	Phe	Thr	Ile	Arg
				65					70					75
Leu	Pro	Lys	Glu	Thr	Pro	Gly	Gln	His	Phe	Glu	Gly	Leu	Asn	Met
				80					85					90
Leu	Thr	Ala	Leu	Leu	Ala	Pro	Arg	Lys	Ser	Arg	Thr	Ala	Lys	Pro
				95					100					105
Leu	Val	Glu	Glu	Ile	Gly	Ala	Ser	Glu	Ile	Pro	Glu	Glu	Val	Val
				110					115					120
Asp	Asp	Glu	Glu	Phe	Asp	Trp	Glu	Ile	Glu	Gln	Thr	Pro	Cys	Glu
				125					130					135
Glu	Val	Ser	Glu	Ser	Ala	Leu	Asn	Pro	Gln	Cys	His	Tyr	Gly	Phe
				140					145					150
Gly	Asn	Leu	Arg	Ser	Gly	Val	Leu	Gln	Arg	Leu	Gln	Asp	Glu	Leu
				155					160					165
Ser	Asp	Val	Ile	Asp	Ile	Lys	Asp	Pro	Asp	Phe	Thr	Pro	Ala	Ala
				170					175					180
Glu	Arg	Arg	Gln	Lys	Arg	Leu	Ala	Ala	Glu	Leu	Ala	Lys	Phe	Asp
				185					190					195
Pro	Asp	His	Tyr	Leu	Ala	Asp	Phe	Phe	Glu	Asp	Glu	Ala	Ile	Glu
				200					205					210
Gln	Ile	Leu	Lys	Tyr	Asn	Pro	Trp	Trp	Thr	Asp	Lys	Tyr	Ser	Lys
				215					220					225
Met	Met	Ala	Phe	Leu	Glu	Lys	Ser	Gln	Glu	Gln	Glu	Asn	His	Ala
				230					235					240
Thr	Leu	Val	Ser	Phe	Ser	Glu	Glu	Glu	Lys	Tyr	Gln	Leu	Arg	Lys
				245					250					255
Phe	Val	Asn	Lys	Ser	Tyr	Leu	Leu	Asp	Lys	Arg	Ala	Cys	Arg	Gln
				260					265					270
Val	Cys	Tyr	Ser	Leu	Ile	Asp	Ile	Leu	Leu	Ala	Tyr	Cys	Tyr	Glu
				275					280					285
Thr	Arg	Val	Thr	Glu	Gly	Glu	Lys	Asn	Val	Glu	Ser	Ala	Trp	Asn
				290					295					300
Ile	Arg	Lys	Leu	Ser	Pro	Thr	Leu	Cys	Trp	Phe	Glu	Thr	Trp	Thr
				305					310					315
Asn	Val	His	Asp	Ile	Met	Val	Ser	Phe	Gly	Arg	Arg	Val	Leu	Cys
				320					325					330
Tyr	Pro	Leu	Tyr	Arg	His	Phe	Lys	Leu	Val	Met	Lys	Ala	Tyr	Arg
				335					340					345
Asp	Thr	Ile	Lys	Ile	Leu	Gln	Leu	Gly	Lys	Ser	Ala	Val	Leu	Lys
				350					355					360
Cys	Leu	Leu	Asp	Ile	His	Lys	Ile	Phe	Gln	Glu	Asn	Asp	Pro	Ala
				365					370					375
Tyr	Ile	Leu	Asn	Asp	Leu	Tyr	Ile	Ser	Asp	Tyr	Cys	Val	Trp	Ile
				380					385					390
Gln	Lys	Val	Lys	Ser	Lys	Lys	Leu	Ala	Ala	Leu	Ala	Glu	Ala	Leu
				395					400					405
Lys	Glu	Val	Ser	Leu	Thr	Lys	Ala	Gln	Leu	Gly	Leu	Glu	Leu	Glu
				410					415					420
Glu	Leu	Glu	Ala	Ala	Ala	Leu	Leu	Val	Gln	Glu	Glu	Glu	Thr	Ala
				425					430					435
Leu	Lys	Ala	Ala	His	Ser	Val	Ser	Gly	Gln	Gln	Thr	Leu	Cys	Ser
				440					445					450
Ser	Ser	Glu	Ala	Ser	Asp	Ser	Glu	Asp	Ser	Asp	Ser	Ser	Val	Ser
				455					460					465
Ser	Gly	Asn	Glu	Asp	Ser	Gly	Ser	Asp	Ser	Glu	Gln	Asp	Glu	Leu
				470					475					480
Lys	Asp	Ser	Pro	Ser	Glu	Thr	Val	Asn	Ser	Leu	Gln	Gly	Pro	Phe
				485					490					495
Leu	Glu	Glu	Ser	Ser	Ala	Phe	Leu	Ile	Val	Asp	Gly	Gly	Val	Arg
				500					505					510

Arg	Asn	Thr	Ala	Ile	Gln	Glu	Ser	Asp	Ala	Ser	Gln	Gly	Lys	Pro
				515					520					525
Leu	Ala	Ser	Ser	Trp	Pro	Leu	Gly	Val	Ser	Gly	Pro	Leu	Ile	Glu
				530					535					540
Glu	Leu	Gly	Glu	Gln	Leu	Lys	Thr	Thr	Val	Gln	Val	Ser	Glu	Pro
				545					550					555
Lys	Gly	Thr	Thr	Ala	Val	Asn	Arg	Ser	Asn	Ile	Gln	Glu	Arg	Asp
				560					565					570
Gly	Cys	Gln	Thr	Pro	Asn	Asn								
				575										

<210> 44  
 <211> 313  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3401751CD1

<400> 44

Met	Ile	Gln	His	Asn	Cys	Ser	Arg	Gln	Gly	Pro	Thr	Ala	Pro	Pro
1				5					10					15
Pro	Pro	Arg	Gly	Pro	Ala	Leu	Pro	Gly	Ala	Gly	Ser	Gly	Leu	Pro
				20					25					30
Ala	Pro	Asp	Pro	Cys	Asp	Tyr	Glu	Gly	Arg	Phe	Ser	Arg	Leu	His
				35					40					45
Gly	Arg	Pro	Pro	Gly	Phe	Leu	His	Cys	Ala	Ser	Phe	Gly	Asp	Pro
				50					55					60
His	Val	Arg	Ser	Phe	His	His	His	Phe	His	Thr	Cys	Arg	Val	Gln
				65					70					75
Gly	Ala	Trp	Pro	Leu	Leu	Asp	Asn	Asp	Phe	Leu	Phe	Val	Gln	Ala
				80					85					90
Thr	Ser	Ser	Pro	Met	Ala	Leu	Gly	Ala	Asn	Ala	Thr	Ala	Thr	Arg
				95					100					105
Lys	Leu	Thr	Ile	Ile	Phe	Lys	Asn	Met	Gln	Glu	Cys	Ile	Asp	Gln
				110					115					120
Lys	Val	Tyr	Gln	Ala	Glu	Val	Asp	Asn	Leu	Pro	Val	Ala	Phe	Glu
				125					130					135
Asp	Gly	Ser	Ile	Asn	Gly	Gly	Asp	Arg	Pro	Gly	Gly	Ser	Ser	Leu
				140					145					150
Ser	Ile	Gln	Thr	Ala	Asn	Pro	Gly	Asn	His	Val	Glu	Ile	Gln	Ala
				155					160					165
Ala	Tyr	Ile	Gly	Thr	Thr	Ile	Ile	Ile	Arg	Gln	Thr	Ala	Gly	Gln
				170					175					180
Leu	Ser	Phe	Ser	Ile	Lys	Val	Ala	Glu	Asp	Val	Ala	Met	Ala	Phe
				185					190					195
Ser	Ala	Glu	Gln	Asp	Leu	Gln	Leu	Cys	Val	Gly	Gly	Cys	Pro	Pro
				200					205					210
Ser	Gln	Arg	Leu	Ser	Arg	Ser	Glu	Arg	Asn	Arg	Arg	Gly	Ala	Ile
				215					220					225
Thr	Ile	Asp	Thr	Ala	Arg	Arg	Leu	Cys	Lys	Glu	Gly	Leu	Pro	Val
				230					235					240
Glu	Asp	Ala	Tyr	Phe	His	Ser	Cys	Val	Phe	Asp	Val	Leu	Ile	Ser
				245					250					255
Gly	Asp	Pro	Asn	Phe	Thr	Val	Ala	Ala	Gln	Ala	Ala	Leu	Glu	Asp
				260					265					270
Ala	Arg	Ala	Phe	Leu	Pro	Asp	Leu	Glu	Lys	Leu	His	Leu	Phe	Pro
				275					280					285
Ser	Asp	Ala	Gly	Val	Pro	Leu	Ser	Ser	Ala	Thr	Leu	Leu	Ala	Pro
				290					295					300
Leu	Leu	Ser	Gly	Leu	Phe	Val	Leu	Trp	Leu	Cys	Ile	Gln		
				305					310					

<210> 45  
 <211> 837  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 045680CD1

<400> 45  
 Met Val Val Ser Ser Asn Gly Ala Ile Met Thr Thr Lys Leu Val  
 1 5 10 15  
 Thr Thr Pro Thr Gly Thr Gln Ala Thr Tyr Thr Arg Pro Thr Val  
 20 25 30  
 Ser Pro Ser Ile Gly Arg Met Ala Ala Thr Pro Gly Ala Ala Thr  
 35 40 45  
 Tyr Val Lys Thr Thr Ser Gly Ser Ile Ile Thr Val Val Pro Lys  
 50 55 60  
 Ser Leu Ala Thr Leu Gly Gly Lys Ile Ile Ser Ser Asn Ile Val  
 65 70 75  
 Ser Gly Thr Thr Thr Lys Ile Thr Thr Ile Pro Met Thr Ser Lys  
 80 85 90  
 Pro Asn Val Ile Val Val Gln Lys Thr Thr Gly Lys Gly Thr Thr  
 95 100 105  
 Ile Gln Gly Leu Pro Gly Lys Asn Val Val Thr Thr Leu Leu Asn  
 110 115 120  
 Ala Gly Gly Glu Lys Thr Ile Gln Thr Val Pro Thr Gly Ala Lys  
 125 130 135  
 Pro Ala Ile Leu Thr Ala Thr Arg Pro Ile Thr Lys Met Ile Val  
 140 145 150  
 Thr Gln Pro Lys Gly Ile Gly Ser Thr Val Gln Pro Ala Ala Lys  
 155 160 165  
 Ile Ile Pro Thr Lys Ile Val Tyr Gly Gln Gln Gly Lys Thr Gln  
 170 175 180  
 Val Leu Ile Lys Pro Lys Pro Val Thr Phe Gln Ala Thr Val Val  
 185 190 195  
 Ser Glu Gln Thr Arg Gln Leu Val Thr Glu Thr Leu Gln Gln Ala  
 200 205 210  
 Ser Arg Val Ala Glu Ala Gly Asn Ser Ser Ile Gln Glu Gly Lys  
 215 220 225  
 Glu Glu Pro Gln Asn Tyr Thr Asp Ser Ser Ser Ser Ser Thr Glu  
 230 235 240  
 Ser Ser Gln Ser Ser Gln Asp Ser Gln Pro Val Val His Val Ile  
 245 250 255  
 Ala Ser Arg Arg Gln Asp Trp Ser Glu His Glu Ile Ala Met Glu  
 260 265 270  
 Thr Ser Pro Thr Ile Ile Tyr Gln Asp Val Ser Ser Glu Ser Gln  
 275 280 285  
 Ser Ala Thr Ser Thr Ile Lys Ala Leu Leu Glu Leu Gln Gln Thr  
 290 295 300  
 Thr Val Lys Glu Lys Leu Glu Ser Lys Pro Arg Gln Pro Thr Ile  
 305 310 315  
 Asp Leu Ser Gln Met Ala Val Pro Ile Gln Met Thr Gln Glu Lys  
 320 325 330  
 Arg His Ser Pro Glu Ser Pro Ser Ile Ala Val Val Glu Ser Glu  
 335 340 345  
 Leu Val Ala Glu Tyr Ile Thr Thr Glu Arg Thr Asp Glu Gly Thr  
 350 355 360  
 Glu Val Ala Phe Pro Leu Leu Val Ser His Arg Ser Gln Pro Gln  
 365 370 375  
 Gln Pro Ser Gln Pro Gln Arg Thr Leu Leu Gln His Val Ala Gln  
 380 385 390  
 Ser Gln Thr Ala Thr Gln Thr Ser Val Val Val Lys Ser Ile Pro

	395		400		405
Ala Ser Ser Pro	Gly Ala Ile Thr His	Ile Met Gln Gln Ala	Leu		
	410		415		420
Ser Ser His Thr	Ala Phe Thr Lys His	Ser Glu Glu Leu Gly	Thr		
	425		430		435
Glu Glu Gly Glu	Val Glu Glu Met Asp	Thr Leu Asp Pro Gln	Thr		
	440		445		450
Gly Leu Phe Tyr	Arg Ser Ala Leu Thr	Gln Ser Gln Ser Ala	Lys		
	455		460		465
Gln Gln Lys Leu	Ser Gln Pro Pro Leu	Glu Gln Thr Gln Leu	Gln		
	470		475		480
Val Lys Thr Leu	Gln Cys Phe Gln Thr	Lys Gln Lys Gln Thr	Ile		
	485		490		495
His Leu Gln Ala	Asp Gln Leu Gln His	Lys Leu Pro Gln Met	Pro		
	500		505		510
Gln Leu Ser Ile	Arg His Gln Lys Leu	Thr Pro Leu Gln Gln	Glu		
	515		520		525
Gln Ala Gln Pro	Lys Pro Asp Val Gln	His Thr Gln His Pro	Met		
	530		535		540
Val Ala Lys Asp	Arg Gln Leu Pro Thr	Leu Met Ala Gln Pro	Pro		
	545		550		555
Gln Thr Val Val	Gln Val Leu Ala Val	Lys Thr Thr Gln Gln	Leu		
	560		565		570
Pro Lys Leu Gln	Gln Ala Pro Asn Gln	Pro Lys Ile Tyr Val	Gln		
	575		580		585
Pro Gln Thr Pro	Gln Ser Gln Met Ser	Leu Pro Ala Ser Ser	Glu		
	590		595		600
Lys Gln Thr Ala	Ser Gln Val Glu Gln	Pro Ile Ile Thr Gln	Gly		
	605		610		615
Ser Ser Val Thr	Lys Ile Thr Phe Glu	Gly Arg Gln Pro Pro	Thr		
	620		625		630
Val Thr Lys Ile	Thr Gly Gly Ser Ser	Val Pro Lys Leu Thr	Ser		
	635		640		645
Pro Val Thr Ser	Ile Ser Pro Ile Gln	Ala Ser Glu Lys Thr	Ala		
	650		655		660
Val Ser Asp Ile	Leu Lys Met Ser Leu	Met Glu Ala Gln Ile	Asp		
	665		670		675
Thr Asn Val Glu	His Met Ile Val Asp	Pro Pro Lys Lys Ala	Leu		
	680		685		690
Ala Thr Ser Met	Leu Thr Gly Glu Ala	Gly Ser Leu Pro Ser	Thr		
	695		700		705
His Met Val Val	Ala Gly Met Ala Asn	Ser Thr Pro Gln Gln	Gln		
	710		715		720
Lys Cys Arg Glu	Ser Cys Ser Ser Pro	Ser Thr Val Gly Ser	Ser		
	725		730		735
Leu Thr Thr Arg	Lys Ile Asp Pro Pro	Ala Val Pro Ala Thr	Gly		
	740		745		750
Gln Phe Met Arg	Ile Gln Asn Val Gly	Gln Lys Lys Ala Glu	Glu		
	755		760		765
Ser Pro Ala Glu	Ile Ile Ile Gln Ala	Ile Pro Gln Tyr Ala	Ile		
	770		775		780
Pro Cys His Ser	Ser Ser Asn Val Val	Val Glu Pro Ser Gly	Leu		
	785		790		795
Leu Glu Leu Asn	Asn Phe Thr Ser Gln	Gln Leu Asp Asp Glu	Glu		
	800		805		810
Thr Ala Met Glu	Gln Asp Ile Asp Ser	Ser Thr Glu Asp Gly	Thr		
	815		820		825
Glu Pro Ser Pro	Ser Gln Ser Ser Ala	Glu Arg Ser			
	830		835		

<210> 46  
 <211> 195  
 <212> PRT



Lys	Gly	Thr	Leu	Pro	Leu	Ser	Arg	Ile	Cys	Ala	Met	Val	Thr	Thr
				140					145					150
Gln	His	Arg	Glu	Asp	Val	Gly	Ile	Gln	Thr	Phe	Asn	Leu	Pro	Leu
				155					160					165
Thr	Pro	Arg	Ile	Glu	Asn	Arg	Lys	Glu	Leu	Arg	Asn	Gln	Ser	Ser
				170					175					180
Gly	Leu	Leu	Asp	Val	Gly	Leu	Arg	Tyr	Arg	Arg	Ser	Pro	Arg	Thr
				185					190					195
Ala	Glu	Gly	Val	Leu	Ala	Ala	Arg	Thr	Val	Ser	Ile	Ser	Val	Gln
				200					205					210
Ile	Ile	Arg	Ala	Cys	Gly	Leu	Gln	Ala	Ala	Ala	Lys	Ala	Leu	Ala
				215					220					225
Glu	Gln	Glu	Pro	Ala	Leu	Gln	Phe	Ser	Ala	Thr	Val	Gly	Val	Asn
				230					235					240
Ala	Ser	Val	Thr	Thr	His	Leu	Ser	Phe	Leu	Pro	Gln	Gly	Glu	Gln
				245					250					255
Arg	Arg	Thr	His	Pro	Val	Ala	Cys	Ser	Phe	Cys	Pro	Glu	Phe	Ser
				260					265					270
His	His	Val	Glu	Phe	Thr	Cys	Asn	Leu	Val	Thr	Gln	His	Cys	Ser
				275					280					285
Gly	Glu	Ala	Cys	Phe	Leu	Ala	Glu	Leu	Leu	Glu	Phe	Ala	Glu	Val
				290					295					300
Ile	Phe	Ala	Val	Tyr	His	Glu	Asn	Thr	Lys	Ser	Ala	Ser	Asp	Ile
				305					310					315
Ile	Ser	Ile	Glu	Ser	Cys	Lys	Glu	Tyr	Leu	Leu	Gly	Val	Val	Lys
				320					325					330
Val	Pro	Thr	Lys	Glu	Leu	Leu	Ile	Lys	Arg	Ser	Gly	Ile	Thr	Gly
				335					340					345
Trp	Tyr	Pro	Ile	Ile	Leu	Pro	Glu	Asp	Gly	Gly	Leu	Pro	His	Gly
				350					355					360
Leu	Glu	Leu	Met	Gln	Lys	Ile	Val	Gly	Gly	Leu	Glu	Leu	Ser	Ile
				365					370					375
Ser	Phe	Thr	His	Arg	Gly	Asp	Arg	Glu	Arg	Val	Leu	Glu	Ala	Ala
				380					385					390
Glu	His	Leu	Gly	Trp	Ser	Phe	Glu	Asn	Ser	Leu	Lys	Asp	Phe	Val
				395					400					405
Arg	Met	Asp	Glu	Gly	Glu	Pro	Ala	Thr	Val	Thr	Ile	Ser	Thr	Pro
				410					415					420
Arg	Leu	Trp	Leu	Pro	Ile	His	Cys	Val	Leu	Leu	Ala	Gly	His	Asn
				425					430					435
His	Ile	His	Lys	Asn	Thr	Tyr	Cys	Tyr	Leu	Arg	Tyr	Lys	Phe	Tyr
				440					445					450
Asp	His	Glu	Ala	Phe	Trp	Thr	Pro	Leu	Lys	Lys	Pro	Lys	Glu	Ser
				455					460					465
Val	Asn	Lys	Lys	Gln	Ile	Met	Val	Thr	Phe	Lys	Ala	Ser	Lys	Arg
				470					475					480
Ala	Glu	Val	Thr	Arg	Gly	Pro	Ser	Leu	Leu	Trp	Tyr	Phe	Arg	Glu
				485					490					495
Glu	Arg	Leu	Glu	Ile	Gln	Val	Trp	Arg	Ala	Tyr	Gly	Asn	Asp	Ser
				500					505					510
Val	Glu	Arg	Pro	His	Gln	Thr	Asp	Ser	Trp	Ile	Gly	Ser	Ala	Tyr
				515					520					525
Val	Asp	Leu	Ala	Arg	Leu	Gly	Glu	Arg	Ser	Ala	Arg	Thr	Leu	Thr
				530					535					540
Val	Ser	Gly	Val	Tyr	Pro	Leu	Phe	Gly	Arg	Asn	Ala	Ser	Asn	Leu
				545					550					555
Ser	Gly	Ala	Ala	Leu	Arg	Val	His	Val	Val	Leu	Ser	Ser	Leu	Ser
				560					565					570
Ser	His	Leu	Glu	Pro	Thr	His	Glu	Leu	Asp	Ser	Met	Asp	Cys	Ser
				575					580					585
Ser	His	Ser	Glu	Ser	Glu	Gln	Leu	Pro	Arg	Arg	Asn	Asp	Glu	Val
				590					595					600
Gln	Leu	Ser	Pro	Pro	Glu	Val	Ile	Ser	Cys	His	Gln	Lys	Ser	Pro



Ala Ser Thr Gln Val	605	Pro Cys Ser Ser	610	Thr Thr Ala Glu Val	615
Leu Thr Gln Glu Gly	620	Pro Ala Asp Leu	625	Asp Gly Thr Phe Ala	630
Ser Ile Leu Val Glu	635	Arg Ala Met His	640	Leu Ser Leu Lys Gly	645
Pro Leu Thr Glu Arg	650	Lys Val Ser Ile	655	Pro Ser Cys Cys Val	660
Phe Ala Thr Ala Asp	665	Glu Ser Ser Pro	670	Val Tyr Thr Gln Val	675
Glu Asn Thr Asp Ser	680	Pro Ile Trp Asn	685	Phe Gln Gln Gln Ser	690
Leu Ser Lys Glu Leu	695	Leu Leu Asp Pro	700	Gln Gln Thr Leu Val	705
Lys Val Trp His Lys	710	Gly Asp Glu Glu	715	Arg Val Ile Gly Phe	720
Ser Val Asp Leu Ser	725	Pro Leu Leu Ser	730	Gly Phe Gln Phe Val	735
Gly Trp Tyr Asn Ile	740	Thr Asp Phe Ser	745	Gly Glu Cys Gln Gly	750
Ile Lys Val Ala Val	755	Ser Pro Leu Glu	760	Ser Leu Ile His Phe	765
Glu Glu Arg Gln Ala	770	Arg Arg Gly Val	775	Glu Thr Ser Lys Ser	780
Ile Pro Ile Tyr Ser	785	Pro Phe Ser Phe	790	Pro Ala Ser Asp Thr	795
Ala Ala Phe Ser Ser	800	His Met Ala Arg	805	Gln Thr Leu Asp Gln	810
Ala His Ala Ser Ser	815	Lys Glu Leu Asp	820	Phe Ser Ser Pro Gly	825
Ser Asp Thr Thr Arg	830	Ser Gln Ala Ser	835	His Glu Glu His Val	840
Gln Asn Ile Arg Arg	845	Phe His Glu Ser	850	Leu His Leu Gln Gly	855
Ala Pro Leu Pro Cys	860	Asp Asp Lys Leu	865	Thr Thr Ser Pro Leu	870
Ser Gln Thr Ser Ile	875	Leu Thr Ser Leu	880	Arg Lys Asn Leu Ser	885
Leu Asp Gln Ile Gln	890	Arg Tyr Phe Arg	895	Gln Lys Leu Thr Lys	900
Phe Leu Pro Leu Ser	905	Pro Gln Thr Gln	910	Thr Ala Ile Ser Gln	915
Gln Glu Ser Cys Arg	920	Asp His Leu Gly	925	Pro Gly Ala Ser Ser	930
Asp Pro Gly Ser Gln	935	Cys Ile Leu Glu	940	Lys Ser Ser Asn Leu	945
Leu Gln Val Ser Ser	950	Leu Ile Thr Asp	955	Gln Gln Thr Ile Thr	960
Asp Ser Gln Ala Ala	965	Leu Ser Ser His	970	Arg Ala Arg Ser Arg	975
Asn Lys Ala Thr Thr	980	Leu Pro Asp Ala	985	Gln Asp Thr Glu Ala	990
Gln Glu Arg Cys Thr	995	Met Pro Asp Glu	1000	Pro Leu Val Arg Ala	1005
Asp Lys Gly Thr Asp	1010	Ser Pro Ser Pro	1015	Pro Pro Leu Glu Glu	1020
Ser Asn Gly Gly Arg	1025	Met Leu His Glu	1030	Ser Leu Arg His Ala	1035
Pro Ile Thr Arg Met	1040	Gln Ser Ser Glu	1045	Thr Leu Arg His Ala	1050
Ala Tyr Ser Asp Glu	1055	Asp Tyr Glu Glu	1060	Asp Thr Glu Ala Gly	1065
	1070		1075	Ile Ile Glu Pro Arg	1080

```

Thr Leu Asn Glu Ile Thr Thr Val Thr Asp Lys Thr Ser Pro Trp
1085 1090 1095
Ser Ser Val Ile Ser Asp Thr Ser Glu Val Ile Ser Pro Gln Pro
1100 1105 1110
Asp Glu Val Gln Arg Glu Gly Pro Ser Cys Pro Ser Pro Gly Pro
1115 1120 1125
Phe Cys Arg Glu Glu Leu Met Val Lys Ser Ser Phe Leu Ser Ser
1130 1135 1140
Pro Glu Arg Ala Val Asn Pro His Leu Pro Arg Gln Gly Ser Pro
1145 1150 1155
Ser Gln Ser Leu Val Ala Cys Glu Cys Glu Ala Ser Lys Ala Arg
1160 1165 1170
Val Gly Gly Glu Ser Ala Ser Ala Asn Pro Gln Pro Ile Pro Cys
1175 1180 1185
Pro Thr Leu Ser Gly Ala Gln Gln Ser Ser Thr Phe Val Gly Trp
1190 1195 1200
Ser Ser Pro Gln Thr Asp Gln Asn Lys Glu Pro Lys Ser Glu Ala
1205 1210 1215
Pro Ala Glu Asn Glu Ala Ala Thr Ser Glu Leu Gly Asp Ser Ala
1220 1225 1230
Asp Ser Phe Lys Lys Leu Pro Leu Asn Leu Ala Ser Gln Ser Arg
1235 1240 1245
Arg Glu Asn His Lys Gly Pro Pro Ile Asp Ser Ser Asp Ile Arg
1250 1255 1260
Gln Arg Gln Val Thr Thr Gly Ser Glu Thr Ser Thr Lys Gln Ser
1265 1270 1275
Leu Leu Leu Pro Gly Pro Ile Val Val Pro Asn Phe Phe Leu Pro
1280 1285 1290
Pro Gln Gln Leu Glu Ala Ser Leu Arg Met Leu Ser Leu Ser Ala
1295 1300 1305
Thr Leu Pro Pro Ala Ala Thr Thr Asp Gln Asp Lys Ser Glu Ala
1310 1315 1320
Thr Arg Gly Ala Leu Ser Gln Arg Pro Cys Arg Pro Arg Pro Asn
1325 1330 1335
Ser Leu Pro Leu Asn Leu Pro Glu Glu Glu Thr Leu Arg Ile Ala
1340 1345 1350
Arg Ile Phe Ser Ser Gln Tyr Ser Gln Lys Asp
1355 1360

```

&lt;210&gt; 48

&lt;211&gt; 552

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3251352CD1

&lt;400&gt; 48

```

Met Leu Ala Leu Lys Ala Ala Phe Gly Glu Ala Leu Phe Thr Ala
1 5 10 15
Gln Asp Glu Ala Leu Leu Leu Arg Arg Leu Thr Leu Ala Ala Gln
20 25 30
His Pro Ala Leu Pro Pro Pro Thr His Leu Phe Tyr Leu His Cys
35 40 45
Val Leu Ser Phe Pro Glu Asn Trp Pro Leu Gly Pro Glu Gly Glu
50 55 60
Glu Ala Ala Pro Leu Leu Leu Gly Pro Gln Leu Cys Arg Gly Leu
65 70 75
Leu Pro Ser Leu Leu His Asp Pro Met Ala Leu Leu Ala Arg Leu
80 85 90
His Leu Leu Cys Leu Leu Cys Ala Glu Glu Glu Glu Glu Lys
95 100 105

```

Gly	Gln	Leu	Pro	Ser	Pro	Arg	His	Tyr	Leu	Glu	Glu	Leu	Leu	Ala
				110					115					120
Gly	Leu	Arg	Gln	Arg	Ala	Ala	Leu	Asp	Gly	Gly	Pro	Arg	Ala	Leu
				125					130					135
Ala	Thr	Leu	Cys	Phe	Gln	Ala	Ser	Tyr	Leu	Val	Ala	Cys	Cys	Leu
				140					145					150
Ala	Gly	Gln	Pro	Thr	Val	Leu	Thr	Pro	Leu	Ile	His	Gly	Leu	Ala
				155					160					165
Gln	Leu	Tyr	Gln	Ala	Arg	Pro	Met	Leu	Ala	Pro	His	Phe	Val	Asp
				170					175					180
Leu	Leu	Asp	Gln	Val	Asp	Ser	Glu	Leu	Arg	Glu	Pro	Leu	Lys	Val
				185					190					195
Val	Leu	Arg	Gln	Val	Val	Val	Ser	Arg	Pro	Gly	Arg	Asp	Glu	Ala
				200					205					210
Leu	Cys	Trp	His	Leu	Gln	Met	Leu	Ala	Lys	Val	Ala	Asp	Gly	Asp
				215					220					225
Ala	Gln	Ser	Ala	Thr	Leu	Asn	Phe	Leu	Gln	Ala	Ala	Ala	Ala	His
				230					235					240
Cys	Thr	Asn	Trp	Asp	Leu	Gln	Gln	Gly	Leu	Leu	Arg	Val	Cys	Arg
				245					250					255
Ala	Leu	Leu	Arg	Ala	Gly	Val	Arg	Gly	Gly	Leu	Val	Asp	Leu	Leu
				260					265					270
Gln	Val	Leu	Ala	Arg	Gln	Leu	Glu	Asp	Pro	Asp	Gly	Arg	Asp	His
				275					280					285
Ala	Arg	Leu	Tyr	Tyr	Ile	Leu	Leu	Ala	His	Leu	Ala	Ala	Pro	Lys
				290					295					300
Leu	Gly	Val	Ala	Leu	Gly	Pro	Ser	Leu	Ala	Ala	Pro	Ala	Leu	Ala
				305					310					315
Ser	Ser	Leu	Val	Ala	Glu	Asn	Gln	Gly	Phe	Val	Ala	Ala	Leu	Met
				320					325					330
Val	Gln	Glu	Ala	Pro	Ala	Leu	Val	Arg	Leu	Ser	Leu	Gly	Ser	His
				335					340					345
Arg	Val	Lys	Gly	Pro	Leu	Pro	Val	Leu	Lys	Leu	Gln	Pro	Glu	Ala
				350					355					360
Leu	Glu	Pro	Ile	Tyr	Ser	Leu	Glu	Leu	Arg	Phe	Arg	Val	Glu	Gly
				365					370					375
Gln	Leu	Tyr	Ala	Pro	Leu	Glu	Ala	Val	His	Val	Pro	Cys	Leu	Cys
				380					385					390
Pro	Gly	Arg	Pro	Ala	Arg	Pro	Leu	Leu	Leu	Pro	Leu	Gln	Pro	Arg
				395					400					405
Cys	Pro	Ala	Pro	Ala	Arg	Leu	Asp	Val	His	Ala	Leu	Tyr	Thr	Thr
				410					415					420
Ser	Thr	Gly	Leu	Thr	Cys	His	Ala	His	Leu	Pro	Pro	Leu	Phe	Val
				425					430					435
Asn	Phe	Ala	Asp	Leu	Phe	Leu	Pro	Phe	Pro	Gln	Pro	Pro	Glu	Gly
				440					445					450
Ala	Gly	Leu	Gly	Phe	Phe	Glu	Glu	Leu	Trp	Asp	Ser	Cys	Leu	Pro
				455					460					465
Glu	Gly	Ala	Glu	Ser	Arg	Val	Trp	Cys	Pro	Leu	Gly	Pro	Gln	Gly
				470					475					480
Leu	Glu	Gly	Leu	Val	Ser	Arg	His	Leu	Glu	Pro	Phe	Val	Val	Val
				485					490					495
Ala	Gln	Pro	Pro	Thr	Ser	Tyr	Cys	Val	Ala	Ile	His	Leu	Pro	Pro
				500					505					510
Asp	Ser	Lys	Leu	Leu	Leu	Arg	Leu	Glu	Ala	Ala	Leu	Ala	Asp	Gly
				515					520					525
Val	Pro	Val	Ala	Leu	Arg	Thr	Asp	Asp	Trp	Ala	Val	Leu	Pro	Leu
				530					535					540
Ala	Gly	Asp	Tyr	Leu	Arg	Gly	Leu	Ala	Ala	Ala	Val			
				545					550					

&lt;210&gt; 49

&lt;211&gt; 345

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 55091643CD1

&lt;400&gt; 49

Met	Glu	Pro	Gln	Lys	Ile	Met	Pro	Pro	Ser	Lys	Pro	His	Pro	Pro	1	5	10	15
Val	Val	Gly	Lys	Val	Thr	His	His	Ser	Ile	Glu	Leu	Tyr	Trp	Asp	20	25	30	35
Leu	Glu	Lys	Lys	Ala	Lys	Arg	Gln	Gly	Pro	Gln	Glu	Gln	Trp	Phe	40	45	50	55
Ser	Phe	Ser	Ile	Glu	Glu	Asp	Pro	Lys	Met	His	Thr	Tyr	Gly		60	65	70	75
Ile	Ile	Tyr	Thr	Gly	Tyr	Ala	Thr	Lys	His	Val	Val	Glu	Gly	Leu	80	85	90	95
Glu	Pro	Arg	Thr	Leu	Tyr	Arg	Phe	Arg	Leu	Lys	Val	Thr	Ser	Pro	100	105	110	115
Ser	Gly	Glu	Cys	Glu	Tyr	Ser	Pro	Leu	Val	Ser	Val	Ser	Thr	Thr	120	125	130	135
Arg	Glu	Pro	Ile	Ser	Ser	Glu	His	Leu	His	Arg	Ala	Val	Ser	Val	140	145	150	155
Asn	Asp	Glu	Asp	Leu	Leu	Val	Arg	Ile	Leu	Gln	Gly	Gly	Arg	Val	160	165	170	175
Lys	Val	Asp	Val	Pro	Asn	Lys	Phe	Gly	Phe	Thr	Ala	Leu	Met	Val	180	185	190	195
Ala	Ala	Gln	Lys	Gly	Tyr	Thr	Arg	Leu	Val	Lys	Ile	Leu	Val	Ser	200	205	210	215
Asn	Gly	Thr	Asp	Val	Asn	Leu	Lys	Asn	Gly	Ser	Gly	Lys	Asp	Ser	220	225	230	235
Leu	Met	Leu	Ala	Cys	Tyr	Ala	Gly	His	Leu	Asp	Val	Val	Lys	Tyr	240	245	250	255
Leu	Arg	Arg	His	Gly	Ala	Ser	Trp	Gln	Ala	Arg	Asp	Leu	Gly	Gly	260	265	270	275
Cys	Thr	Ala	Leu	His	Trp	Ala	Ala	Asp	Gly	Gly	His	Cys	Ser	Val	280	285	290	295
Ile	Glu	Trp	Met	Ile	Lys	Asp	Gly	Cys	Glu	Val	Asp	Val	Val	Asp	300	305	310	315
Thr	Gly	Ser	Gly	Trp	Thr	Pro	Leu	Met	Arg	Val	Ser	Ala	Val	Ser	320	325	330	335
Gly	Asn	Gln	Arg	Val	Ala	Ser	Leu	Leu	Ile	Asp	Ala	Gly	Ala	Asn	340	345	350	355
Val	Asn	Val	Lys	Asp	Arg	Asn	Gly	Lys	Thr	Pro	Leu	Met	Val	Ala	360	365	370	375
Val	Leu	Asn	Asn	His	Glu	Glu	Leu	Val	Gln	Leu	Leu	Leu	Asp	Lys	380	385	390	395
Gly	Ala	Asp	Ala	Ser	Val	Lys	Asn	Glu	Phe	Gly	Lys	Gly	Val	Leu	400	405	410	415
Glu	Met	Ala	Arg	Val	Phe	Asp	Arg	Gln	Ser	Val	Val	Ser	Leu	Leu	420	425	430	435
Glu	Glu	Arg	Lys	Lys	Lys	Gln	Arg	Pro	Lys	Lys	Ser	Cys	Val	Cys	440	445	450	455

&lt;210&gt; 50

&lt;211&gt; 30

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

<223> Incyte ID No: 7500770CD1

<400> 50

Met	Ser	Cys	Pro	Val	Gln	Thr	Met	Asp	Pro	Glu	Val	Thr	Leu	Leu
1				5					10					15
Leu	Gln	Pro	Pro	Ala	Pro	Gly	Pro	Arg	Trp	Ala	Ala	Gly	Gly	Thr
				20					25					30

<210> 51

<211> 685

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7506350CD1

<400> 51

Met	Ala	Leu	Gly	Glu	Glu	Lys	Ala	Glu	Ala	Glu	Ala	Ser	Glu	Asp
1				5					10					15
Thr	Lys	Ala	Gln	Ser	Tyr	Gly	Arg	Gly	Ser	Cys	Arg	Glu	Arg	Glu
				20					25					30
Leu	Asp	Ile	Pro	Gly	Pro	Met	Ser	Gly	Glu	Gln	Pro	Pro	Arg	Leu
				35					40					45
Glu	Ala	Glu	Gly	Gly	Leu	Ile	Ser	Pro	Val	Trp	Gly	Ala	Glu	Gly
				50					55					60
Ile	Pro	Ala	Pro	Thr	Cys	Trp	Ile	Gly	Thr	Asp	Pro	Gly	Gly	Pro
				65					70					75
Ser	Arg	Ala	His	Gln	Pro	Gln	Ala	Ser	Asp	Ala	Asn	Arg	Glu	Pro
				80					85					90
Val	Ala	Glu	Arg	Ser	Glu	Pro	Ala	Leu	Ser	Gly	Leu	Pro	Pro	Ala
				95					100					105
Thr	Met	Gly	Ser	Gly	Asp	Leu	Leu	Leu	Ser	Gly	Glu	Ser	Gln	Val
				110					115					120
Glu	Lys	Thr	Lys	Leu	Ser	Ser	Ser	Glu	Glu	Phe	Pro	Gln	Thr	Leu
				125					130					135
Ser	Leu	Pro	Arg	Thr	Thr	Thr	Ile	Cys	Ser	Gly	His	Asp	Ala	Asp
				140					145					150
Thr	Glu	Asp	Asp	Pro	Ser	Leu	Ala	Asp	Leu	Pro	Gln	Ala	Leu	Asp
				155					160					165
Leu	Ser	Gln	Gln	Pro	His	Ser	Ser	Gly	Leu	Ser	Cys	Leu	Ser	Gln
				170					175					180
Trp	Lys	Ser	Val	Leu	Ser	Pro	Gly	Ser	Ala	Ala	Gln	Pro	Ser	Ser
				185					190					195
Cys	Ser	Ile	Ser	Ala	Ser	Ser	Thr	Gly	Ser	Ser	Leu	Gln	Gly	His
				200					205					210
Gln	Glu	Arg	Ala	Glu	Pro	Arg	Gly	Gly	Ser	Leu	Ala	Lys	Val	Ser
				215					220					225
Ser	Ser	Leu	Glu	Pro	Val	Val	Pro	Gln	Glu	Pro	Ser	Ser	Val	Val
				230					235					240
Gly	Leu	Gly	Pro	Arg	Pro	Gln	Trp	Ser	Pro	Gln	Pro	Val	Phe	Ser
				245					250					255
Gly	Gly	Asp	Ala	Ser	Gly	Leu	Gly	Arg	Arg	Arg	Leu	Ser	Phe	Gln
				260					265					270
Ala	Glu	Tyr	Trp	Ala	Cys	Val	Leu	Pro	Asp	Ser	Leu	Pro	Pro	Ser
				275					280					285
Pro	Asp	Arg	His	Ser	Pro	Leu	Trp	Asn	Pro	Asn	Lys	Glu	Tyr	Glu
				290					295					300
Asp	Leu	Leu	Asp	Tyr	Thr	Tyr	Pro	Leu	Arg	Pro	Gly	Pro	Gln	Leu
				305					310					315
Pro	Lys	His	Leu	Asp	Ser	Arg	Val	Pro	Ala	Asp	Pro	Val	Leu	Gln
				320					325					330

Asp Ser Gly Val	Asp Leu Asp Ser Phe	Ser Val Ser Pro Ala	Ser
335	340		345
Thr Leu Lys Ser	Pro Thr Asn Val Ser	Pro Asn Cys Pro Pro	Ala
350	355		360
Glu Ala Thr Ala	Leu Pro Phe Ser Gly	Pro Arg Glu Pro Ser	Leu
365	370		375
Lys Gln Trp Pro	Ser Arg Val Pro Gln	Lys Gln Gly Gly Met	Gly
380	385		390
Leu Ala Ser Trp	Ser Gln Leu Ala Ser	Thr Pro Arg Ala Pro	Gly
395	400		405
Ser Arg Asp Ala	Arg Trp Glu Arg Arg	Glu Pro Ala Leu Arg	Gly
410	415		420
Ala Lys Asp Arg	Leu Thr Ile Gly Lys	His Leu Asp Met Gly	Ser
425	430		435
Pro Gln Leu Arg	Thr Arg Asp Arg Gly	Trp Pro Ser Pro Arg	Pro
440	445		450
Glu Arg Glu Lys	Arg Thr Ser Gln Ser	Ala Arg Arg Pro Thr	Cys
455	460		465
Thr Glu Ser Arg	Trp Lys Ser Glu Glu	Glu Val Glu Ser Asp	Asp
470	475		480
Glu Tyr Leu Ala	Leu Pro Ala Arg Leu	Thr Gln Val Ser Ser	Leu
485	490		495
Val Ser Tyr Leu	Gly Ser Ile Ser Thr	Leu Val Thr Leu Pro	Thr
500	505		510
Gly Asp Ile Lys	Gly Gln Ser Pro Leu	Glu Val Ser Asp Ser	Asp
515	520		525
Gly Pro Ala Ser	Phe Pro Ser Ser Ser	Ser Gln Ser Gln Leu	Pro
530	535		540
Arg Lys Gly Gly	Glu Gln Gly Lys Glu	Ser Leu Val Gln Cys	Val
545	550		555
Lys Thr Phe Cys	Cys Gln Leu Glu Glu	Leu Ile Cys Trp Leu	Tyr
560	565		570
Asn Val Ala Asp	Val Thr Asp His Gly	Thr Ala Ala Arg Ser	Asn
575	580		585
Leu Thr Ser Leu	Lys Ser Ser Leu Gln	Leu Tyr Arg Gln Phe	Lys
590	595		600
Lys Asp Ile Asp	Glu His Gln Ser Leu	Thr Glu Ser Val Leu	Gln
605	610		615
Lys Gly Glu Ile	Leu Leu Gln Cys Leu	Leu Glu Asn Thr Pro	Val
620	625		630
Leu Glu Asp Val	Leu Gly Arg Ile Ala	Lys Gln Ser Gly Glu	Leu
635	640		645
Glu Ser His Ala	Asp Arg Leu Tyr Asp	Ser Ile Leu Ala Ser	Leu
650	655		660
Asp Met Leu Ala	Gly Cys Thr Leu Ile	Pro Asp Lys Lys Pro	Met
665	670		675
Ala Ala Met Glu	His Pro Cys Glu Gly	Val	
680	685		

&lt;210&gt; 52

&lt;211&gt; 104

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7508370CD1

&lt;400&gt; 52

Met Ala Ala Val	Asp Asp Leu Gln Phe	Glu Glu Phe Gly Asn	Ala
1	5	10	15
Ala Thr Ser Leu	Thr Ala Asn Pro Asp	Ala Thr Thr Val Asn	Ile
20	25	30	

Glu	Asp	Pro	Gly	Glu	Thr	Pro	Lys	His	Gln	Pro	Gly	Ser	Pro	Arg
				35					40					45
Gly	Ser	Gly	Arg	Glu	Glu	Asp	Asp	Glu	Leu	Leu	Gly	Asn	Asp	Asp
				50					55					60
Ser	Asp	Lys	Thr	Glu	Leu	Leu	Ala	Gly	Gln	Lys	Lys	Ser	Ser	Pro
				65					70					75
Phe	Trp	Thr	Phe	Glu	Tyr	Tyr	Gln	Thr	Phe	Phe	Asp	Val	Asp	Thr
				80					85					90
Tyr	Gln	Ala	Pro	Phe	Gly	Tyr	Val	Pro	Arg	Trp	Ser	Leu	Pro	
				95					100					

&lt;210&gt; 53

&lt;211&gt; 672

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2894093CD1

&lt;400&gt; 53

Met	Val	Ser	Trp	Gln	Asp	Ile	Leu	Gln	Gln	Ile	Asn	Glu	Ile	Asn
1				5					10					15
Thr	Leu	Val	Gly	Ser	Ala	Ser	Ser	Lys	Lys	Ala	Lys	Lys	Pro	Val
				20					25					30
Gly	Gly	Asn	Ala	Pro	Leu	Tyr	Tyr	Glu	Ala	Pro	Gln	Pro	Ala	Val
				35					40					45
Pro	Ala	Pro	Gly	Lys	Lys	Lys	Ala	Gln	Tyr	Glu	Glu	Pro	Gln	Ala
				50					55					60
Pro	Pro	Pro	Val	Thr	Ser	Val	Ile	Thr	Thr	Glu	Val	Asp	Met	Arg
				65					70					75
Tyr	Tyr	Asn	Tyr	Leu	Leu	Asn	Pro	Ile	Arg	Glu	Glu	Phe	Ile	Ser
				80					85					90
Val	Pro	Leu	Ile	Leu	His	Cys	Met	Leu	Glu	Gln	Val	Val	Ala	Thr
				95					100					105
Glu	Glu	Asp	Leu	Val	Pro	Pro	Ser	Leu	Arg	Glu	Pro	Ser	Pro	Arg
				110					115					120
Ala	Asp	Gly	Leu	Asp	His	Arg	Ile	Ala	Ala	His	Ile	Val	Ser	Leu
				125					130					135
Leu	Pro	Ser	Leu	Cys	Leu	Ser	Glu	Arg	Glu	Lys	Lys	Asn	Leu	His
				140					145					150
Asp	Ile	Phe	Leu	Ser	Glu	Glu	Glu	Asn	Glu	Ser	Lys	Ala	Val	Pro
				155					160					165
Lys	Gly	Pro	Leu	Leu	Leu	Asn	Tyr	His	Asp	Ala	His	Ala	His	Lys
				170					175					180
Lys	Tyr	Ala	Leu	Gln	Asp	Gln	Lys	Asn	Phe	Asp	Pro	Val	Gln	Ile
				185					190					195
Glu	Gln	Glu	Met	Gln	Ser	Lys	Leu	Pro	Leu	Trp	Glu	Phe	Leu	Gln
				200					205					210
Phe	Pro	Leu	Pro	Pro	Pro	Trp	Asn	Asn	Thr	Lys	Arg	Leu	Ala	Thr
				215					220					225
Ile	His	Glu	Leu	Met	His	Phe	Cys	Thr	Ser	Asp	Val	Leu	Ser	Trp
				230					235					240
Asn	Glu	Val	Glu	Arg	Ala	Phe	Lys	Val	Phe	Thr	Phe	Glu	Ser	Leu
				245					250					255
Lys	Leu	Ser	Glu	Val	Asp	Glu	Lys	Gly	Lys	Leu	Lys	Pro	Ser	Gly
				260					265					270
Met	Met	Cys	Gly	Ser	Asp	Ser	Glu	Met	Phe	Asn	Ile	Pro	Trp	Asp
				275					280					285
Asn	Pro	Ala	Arg	Phe	Ala	Lys	Gln	Ile	Arg	Gln	Gln	Tyr	Val	Met
				290					295					300
Lys	Met	Asn	Thr	Gln	Glu	Ala	Lys	Gln	Lys	Ala	Asp	Ile	Lys	Ile
				305					310					315

Lys	Asp	Arg	Thr	Leu	Phe	Val	Asp	Gln	Asn	Leu	Ser	Met	Ser	Val
				320					325					330
Gln	Asp	Asn	Glu	Ser	Asn	Arg	Glu	Pro	Ser	Asp	Pro	Ser	Gln	Cys
				335					340					345
Asp	Ala	Asn	Asn	Met	Lys	His	Ser	Asp	Leu	Asn	Asn	Leu	Lys	Leu
				350					355					360
Ser	Val	Pro	Asp	Asn	Arg	Gln	Leu	Leu	Glu	Gln	Glu	Ser	Ile	Met
				365					370					375
Lys	Ala	Gln	Pro	Gln	His	Glu	Ser	Leu	Glu	Gln	Thr	Thr	Asn	Asn
				380					385					390
Glu	Ile	Lys	Asp	Asp	Ala	Val	Thr	Lys	Ala	Asp	Ser	His	Glu	Lys
				395					400					405
Lys	Pro	Lys	Lys	Met	Met	Val	Glu	Ala	Asp	Leu	Glu	Asp	Ile	Lys
				410					415					420
Lys	Thr	Gln	Gln	Arg	Ser	Leu	Met	Asp	Trp	Ser	Phe	Thr	Glu	His
				425					430					435
Phe	Lys	Pro	Lys	Val	Leu	Leu	Gln	Val	Leu	Gln	Glu	Ala	His	Lys
				440					445					450
Gln	Tyr	Arg	Cys	Val	Asp	Ser	Tyr	Tyr	His	Thr	Gln	Asp	Asn	Ser
				455					460					465
Leu	Leu	Leu	Val	Phe	His	Asn	Pro	Met	Asn	Arg	Gln	Arg	Leu	His
				470					475					480
Cys	Glu	Tyr	Trp	Asn	Ile	Ala	Leu	His	Ser	Asn	Val	Gly	Phe	Arg
				485					490					495
Asn	Tyr	Leu	Glu	Leu	Val	Ala	Lys	Ser	Ile	Gln	Asp	Trp	Ile	Thr
				500					505					510
Lys	Glu	Glu	Ala	Ile	Tyr	Gln	Glu	Ser	Lys	Met	Asn	Glu	Lys	Ile
				515					520					525
Ile	Arg	Thr	Arg	Ala	Glu	Leu	Glu	Leu	Lys	Ser	Ser	Ala	Asn	Ala
				530					535					540
Lys	Leu	Thr	Ser	Ala	Ser	Lys	Ile	Phe	Ser	Ile	Lys	Glu	Ser	Lys
				545					550					555
Ser	Asn	Lys	Gly	Ile	Ser	Lys	Thr	Glu	Ile	Ser	Asp	Gln	Glu	Lys
				560					565					570
Glu	Lys	Glu	Lys	Glu	Lys	Ile	Pro	Phe	Ile	Leu	Glu	Gly	Ser	Leu
				575					580					585
Lys	Ala	Trp	Lys	Glu	Glu	Gln	His	Arg	Leu	Ala	Glu	Glu	Glu	Arg
				590					595					600
Leu	Arg	Glu	Glu	Lys	Lys	Ala	Glu	Lys	Lys	Gly	Lys	Glu	Ala	Gly
				605					610					615
Lys	Lys	Lys	Gly	Lys	Asp	Asn	Ala	Glu	Lys	Glu	Asp	Ser	Arg	Ser
				620					625					630
Leu	Lys	Lys	Lys	Ser	Pro	Tyr	Lys	Glu	Lys	Ser	Lys	Glu	Glu	Gln
				635					640					645
Val	Lys	Ile	Gln	Glu	Val	Thr	Glu	Glu	Ser	Pro	His	Gln	Pro	Glu
				650					655					660
Pro	Lys	Ile	Thr	Tyr	Pro	Ala	Gln	Leu	Leu	Ser	Lys			
				665					670					

&lt;210&gt; 54

&lt;211&gt; 689

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7507335CD1

&lt;400&gt; 54

Met	Ala	Val	Val	Ser	Ala	Val	Arg	Trp	Leu	Gly	Leu	Arg	Ser	Arg
1				5					10					15
Leu	Gly	Gln	Pro	Leu	Thr	Gly	Arg	Arg	Ala	Gly	Leu	Cys	Glu	Gln
				20					25					30



Ala	Arg	Ser	Cys	Arg	Phe	Tyr	Ser	Gly	Ser	Ala	Thr	Leu	Ser	Lys
				35					40					45
Val	Glu	Gly	Thr	Asp	Val	Thr	Gly	Ile	Glu	Glu	Val	Val	Ile	Pro
				50					55					60
Lys	Lys	Lys	Thr	Trp	Asp	Lys	Val	Ala	Val	Leu	Gln	Ala	Leu	Ala
				65					70					75
Ser	Thr	Val	Asn	Arg	Asp	Thr	Thr	Ala	Val	Pro	Tyr	Val	Phe	Gln
				80					85					90
Asp	Asp	Pro	Tyr	Leu	Met	Pro	Ala	Ser	Ser	Leu	Glu	Ser	Arg	Ser
				95					100					105
Phe	Leu	Leu	Ala	Lys	Lys	Ser	Gly	Glu	Asn	Val	Ala	Lys	Phe	Ile
				110					115					120
Ile	Asn	Ser	Tyr	Pro	Lys	Tyr	Phe	Gln	Lys	Asp	Ile	Ala	Glu	Pro
				125					130					135
His	Ile	Pro	Cys	Leu	Met	Pro	Glu	Tyr	Phe	Glu	Pro	Gln	Ile	Lys
				140					145					150
Asp	Ile	Ser	Glu	Ala	Ala	Leu	Lys	Glu	Arg	Ile	Glu	Leu	Arg	Lys
				155					160					165
Val	Lys	Ala	Ser	Val	Asp	Met	Phe	Asp	Gln	Leu	Leu	Gln	Ala	Gly
				170					175					180
Thr	Thr	Val	Ser	Leu	Glu	Thr	Thr	Asn	Ser	Leu	Leu	Asp	Leu	Leu
				185					190					195
Cys	Tyr	Tyr	Gly	Asp	Gln	Glu	Pro	Ser	Thr	Asp	Tyr	His	Phe	Gln
				200					205					210
Gln	Thr	Gly	Gln	Ser	Glu	Ala	Leu	Glu	Glu	Glu	Asn	Asp	Glu	Thr
				215					220					225
Ser	Arg	Arg	Lys	Ala	Gly	His	Gln	Phe	Gly	Val	Thr	Trp	Arg	Ala
				230					235					240
Lys	Asn	Asn	Ala	Glu	Arg	Ile	Phe	Ser	Leu	Met	Pro	Glu	Lys	Asn
				245					250					255
Glu	His	Ser	Tyr	Cys	Thr	Met	Ile	Arg	Gly	Met	Val	Lys	His	Arg
				260					265					270
Ala	Tyr	Glu	Gln	Ala	Leu	Asn	Leu	Tyr	Thr	Glu	Leu	Leu	Asn	Asn
				275					280					285
Arg	Leu	His	Ala	Asp	Val	Tyr	Thr	Phe	Asn	Ala	Leu	Ile	Glu	Ala
				290					295					300
Thr	Val	Cys	Ala	Ile	Asn	Glu	Lys	Phe	Glu	Glu	Lys	Trp	Ser	Lys
				305					310					315
Ile	Leu	Glu	Leu	Leu	Arg	His	Met	Val	Ala	Gln	Lys	Val	Lys	Pro
				320					325					330
Asn	Leu	Gln	Thr	Phe	Asn	Thr	Ile	Leu	Lys	Cys	Leu	Arg	Arg	Phe
				335					340					345
His	Val	Phe	Ala	Arg	Ser	Pro	Ala	Leu	Gln	Val	Leu	Arg	Glu	Met
				350					355					360
Lys	Ala	Ile	Gly	Ile	Glu	Pro	Ser	Leu	Ala	Thr	Tyr	His	His	Ile
				365					370					375
Ile	Arg	Leu	Phe	Asp	Gln	Pro	Gly	Asp	Pro	Leu	Lys	Arg	Ser	Ser
				380					385					390
Phe	Ile	Ile	Tyr	Asp	Ile	Met	Asn	Glu	Leu	Met	Gly	Lys	Arg	Phe
				395					400					405
Ser	Pro	Lys	Asp	Pro	Asp	Asp	Asp	Lys	Phe	Phe	Gln	Ser	Ala	Met
				410					415					420
Ser	Ile	Cys	Ser	Ser	Leu	Arg	Asp	Leu	Glu	Leu	Ala	Tyr	Gln	Val
				425					430					435
His	Gly	Leu	Leu	Lys	Thr	Gly	Asp	Asn	Trp	Lys	Phe	Ile	Gly	Pro
				440					445					450
Asp	Gln	His	Arg	Asn	Phe	Tyr	Tyr	Ser	Lys	Phe	Phe	Asp	Leu	Ile
				455					460					465
Cys	Leu	Met	Glu	Gln	Ile	Asp	Val	Thr	Leu	Lys	Trp	Tyr	Glu	Asp
				470					475					480
Leu	Ile	Pro	Ser	Ala	Tyr	Phe	Pro	His	Ser	Gln	Thr	Met	Ile	His
				485					490					495
Leu	Leu	Gln	Ala	Leu	Asp	Val	Ala	Asn	Arg	Leu	Glu	Val	Ile	Pro

	500		505		510
Lys Ile Trp Lys	Asp Ser Lys Glu Tyr	Gly His Thr Phe Arg	Ser		
	515		520		525
Asp Leu Arg Glu	Glu Ile Leu Met Leu	Met Ala Arg Asp Lys	His		
	530		535		540
Pro Pro Glu Leu	Gln Val Ala Phe Ala	Asp Cys Ala Ala Asp	Ile		
	545		550		555
Lys Ser Ala Tyr	Glu Ser Gln Pro Ile	Arg Gln Thr Ala Gln	Asp		
	560		565		570
Trp Pro Ala Thr	Ser Leu Asn Cys Ile	Ala Ile Leu Phe Leu	Arg		
	575		580		585
Ala Gly Arg Thr	Gln Glu Ala Trp Lys	Met Leu Gly Leu Phe	Arg		
	590		595		600
Lys His Asn Lys	Ile Pro Arg Ser Glu	Leu Leu Asn Glu Leu	Met		
	605		610		615
Asp Ser Ala Lys	Val Ser Asn Ser Pro	Ser Gln Ala Ile Glu	Val		
	620		625		630
Val Glu Leu Ala	Ser Ala Phe Ser Leu	Pro Ile Cys Glu Gly	Leu		
	635		640		645
Thr Gln Arg Val	Met Ser Asp Phe Ala	Ile Asn Gln Glu Gln	Lys		
	650		655		660
Glu Ala Leu Ser	Asn Leu Thr Ala Leu	Thr Ser Asp Ser Asp	Thr		
	665		670		675
Asp Ser Ser Ser	Asp Ser Asp Ser Asp	Thr Ser Glu Gly Lys			
	680		685		

&lt;210&gt; 55

&lt;211&gt; 359

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7509081CD1

&lt;400&gt; 55

Met Ala Ala Pro Cys	Ala Glu Asp Pro Ser	Leu Glu Arg His Phe
1	5	10
Lys Gly His Arg Asp	Ala Val Thr Cys Val	Asp Phe Ser Ile Asn
	20	25
Thr Lys Gln Leu Ala	Ser Gly Ser Met Asp	Ser Cys Leu Met Val
	35	40
Trp His Met Lys Pro	Gln Ser Arg Ala Tyr	Arg Phe Thr Gly His
	50	55
Lys Asp Ala Val Thr	Cys Val Asn Phe Ser	Pro Ser Gly His Leu
	65	70
Leu Ala Ser Gly Ser	Arg Asp Lys Thr Val	Arg Ile Trp Val Pro
	80	85
Asn Val Lys Gly Glu	Ser Thr Val Phe Arg	Ala His Thr Ala Thr
	95	100
Val Arg Ser Val His	Phe Cys Ser Asp Gly	Gln Ser Phe Val Thr
	110	115
Ala Ser Asp Asp Lys	Thr Val Lys Val Trp	Ala Thr His Arg Gln
	125	130
Lys Phe Leu Phe Ser	Leu Ser Gln His Ile	Asn Trp Val Arg Cys
	140	145
Ala Lys Phe Ser Pro	Asp Gly Arg Leu Ile	Val Ser Ala Ser Asp
	155	160
Asp Lys Thr Val Lys	Leu Trp Asp Lys Ser	Ser Arg Glu Cys Val
	170	175
His Ser Tyr Cys Glu	His Gly Gly Phe Val	Thr Tyr Val Asp Phe
	185	190
His Pro Ser Gly Thr	Cys Ile Ala Ala Ala	Gly Met Asp Asn Thr

	200		205		210									
Val	Lys	Val	Trp	Asp	Val	Arg	Thr	His	Arg	Leu	Leu	Gln	His	Tyr
	215								220					225
Gln	Leu	His	Ser	Ala	Ala	Val	Asn	Gly	Leu	Ser	Phe	His	Pro	Ser
	230								235					240
Gly	Asn	Tyr	Leu	Ile	Thr	Ala	Ser	Ser	Asp	Ser	Thr	Leu	Lys	Ile
	245								250					255
Leu	Asp	Leu	Met	Glu	Gly	Arg	Leu	Leu	Tyr	Thr	Leu	His	Gly	His
	260								265					270
Gln	Gly	Pro	Ala	Thr	Thr	Val	Ala	Phe	Ser	Arg	Thr	Gly	Glu	Tyr
	275								280					285
Phe	Ala	Ser	Gly	Gly	Ser	Asp	Glu	Gln	Val	Met	Val	Trp	Lys	Ser
	290								295					300
Asn	Phe	Asp	Ile	Val	Asp	His	Gly	Glu	Val	Thr	Lys	Val	Pro	Arg
	305								310					315
Pro	Pro	Ala	Thr	Leu	Ala	Ser	Ser	Met	Gly	Asn	Leu	Thr	Val	Ser
	320								325					330
Ile	Leu	Glu	Gln	Arg	Leu	Thr	Leu	Thr	Glu	Asp	Lys	Leu	Lys	Gln
	335								340					345
Cys	Leu	Glu	Asn	Gln	Gln	Leu	Ile	Met	Gln	Arg	Ala	Thr	Pro	
	350								355					

&lt;210&gt; 56

&lt;211&gt; 316

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7502450CD1

&lt;400&gt; 56

Met	Thr	Pro	Ser	Glu	Gly	Ala	Arg	Ala	Gly	Thr	Gly	Arg	Glu	Leu
1				5					10					15
Glu	Met	Leu	Asp	Ser	Leu	Leu	Ala	Leu	Gly	Gly	Leu	Val	Met	His
				20					25					30
Gly	Asp	Ser	Val	Glu	Trp	Glu	Gly	Arg	Ser	Leu	Leu	Lys	Ala	Leu
				35					40					45
Val	Lys	Lys	Ser	Ala	Leu	Cys	Gly	Glu	Gln	Val	His	Ile	Leu	Gly
				50					55					60
Cys	Glu	Val	Ser	Glu	Glu	Glu	Phe	Arg	Glu	Gly	Phe	Asp	Ser	Asp
				65					70					75
Ile	Asn	Asn	Arg	Leu	Val	Tyr	His	Asp	Phe	Phe	Arg	Asp	Pro	Leu
				80					85					90
Asn	Trp	Ser	Lys	Thr	Glu	Glu	Ala	Phe	Pro	Gly	Gly	Pro	Leu	Gly
				95					100					105
Ala	Leu	Arg	Ala	Met	Cys	Lys	Arg	Thr	Asp	Pro	Val	Pro	Val	Thr
				110					115					120
Ile	Ala	Leu	Asp	Ser	Leu	Ser	Trp	Leu	Leu	Leu	Arg	Leu	Pro	Cys
				125					130					135
Thr	Thr	Leu	Cys	Gln	Val	Leu	His	Ala	Val	Ser	His	Gln	Asp	Ser
				140					145					150
Cys	Pro	Gly	Asp	Ser	Ser	Ser	Val	Gly	Lys	Val	Ser	Val	Leu	Gly
				155					160					165
Leu	Leu	His	Glu	Glu	Leu	His	Gly	Pro	Gly	Pro	Val	Gly	Ala	Leu
				170					175					180
Ser	Ser	Leu	Ala	Gln	Thr	Glu	Val	Thr	Leu	Gly	Gly	Thr	Met	Gly
				185					190					195
Gln	Ala	Ser	Ala	His	Ile	Leu	Cys	Arg	Arg	Pro	Arg	Gln	Arg	Pro
				200					205					210
Thr	Asp	Gln	Thr	Gln	Trp	Phe	Ser	Ile	Leu	Pro	Asp	Phe	Ser	Leu
				215					220					225
Asp	Leu	Gln	Glu	Gly	Pro	Ser	Val	Glu	Ser	Gln	Pro	Tyr	Ser	Asp

	230		235		240
Pro His Ile Pro	Pro Val Asp Pro Thr	Thr His Leu Thr Phe	Asn		
	245		250		255
Leu His Leu Ser	Lys Glu Arg Glu	Ala Arg Asp Ser Leu	Ile		
	260		265		270
Leu Pro Phe Gln	Phe Ser Ser Glu Lys	Gln Gln Ala Leu Leu	Arg		
	275		280		285
Pro Arg Pro Gly	Gln Ala Thr Ser His	Ile Phe Tyr Glu Pro	Asp		
	290		295		300
Ala Tyr Asp Asp	Leu Asp Gln Glu Asp	Pro Asp Asp Asp Leu	Asp		
	305		310		315
Ile					

<210> 57  
 <211> 60  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7501405CD1

<400> 57	
Met Pro Gly Pro Arg	Val Trp Gly Lys Tyr Leu Trp Arg Ser Pro
1	5 10 15
His Ser Lys Gly Cys	Pro Gly Ala Ile Ser His Ser Ser Val Arg
	20 25 30
Ile Gln Ala Ser Ser	Val Val Ser Lys Ala Pro Leu Trp Ala Gly
	35 40 45
Pro Leu Val Arg Gly	Ser Leu Tyr Pro Gln Gly Gly Val Cys Gly
	50 55 60

<210> 58  
 <211> 205  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7504528CD1

<400> 58	
Met Ala Ala Ala Leu	Gln Val Leu Pro Arg Leu Ala Arg Ala Pro
1	5 10 15
Leu His Pro Leu Leu	Trp Arg Gly Ser Val Ala Arg Leu Ala Ser
	20 25 30
Ser Met Ala Leu Ala	Glu Gln Ala Arg Gln Leu Phe Glu Ser Ala
	35 40 45
Val Gly Ala Val Leu	Pro Gly Pro Met Leu His Arg Ala Leu Ser
	50 55 60
Leu Asp Pro Gly Gly	Arg Gln Leu Lys Val Arg Asp Arg Asn Phe
	65 70 75
Gln Leu Arg Gln Asn	Leu Tyr Leu Val Gly Phe Gly Lys Ala Val
	80 85 90
Leu Gly Met Ala Ala	Ala Ala Glu Glu Leu Leu Gly Gln His Leu
	95 100 105
Val Gln Gly Val Ile	Ser Val Pro Lys Gly Ile Arg Ala Ala Met
	110 115 120
Glu Arg Ala Gly Lys	Gln Glu Met Leu Leu Lys Pro His Ser Arg
	125 130 135
Val Gln Val Phe Glu	Gly Ala Glu Asp Asn Leu Pro Asp Arg Asp

	140		145		150
Ala Leu Arg Ala	Ala Leu Ala Ile Gln	Gln Leu Ala Glu Gly	Leu		
	155		160		165
Thr Ala Asp Asp	Leu Leu Val Leu	Ile Ser Gly Trp Gly	Thr		
	170		175		180
Pro Ala Ala His	Arg Asp Asp Arg Tyr	Gln Cys His Gly His	Pro		
	185		190		195
Pro Leu Val Pro	Ala Ala Ser Val Met	Ala			
	200		205		

<210> 59  
 <211> 203  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7509049CD1

<400> 59  
 Met Leu Gly Ile Thr Ser Cys Ser Asp Gln Gln Ala Lys Glu Gly  
   1                  5                  10                  15  
 Glu Gly Leu Glu Gly Ser Ser Thr Gly Ser Ser Ser Gly Asn His  
                   20                  25                  30  
 Gly Gly Ser Gly Gly Gly Asn Gly His Lys Pro Gly Cys Glu Lys  
                   35                  40                  45  
 Pro Gly Asn Glu Ala Arg Gly Ser Gly Glu Ser Gly Ile Gln Asn  
                   50                  55                  60  
 Ser Glu Thr Ser Pro Gly Met Phe Asn Phe Asp Thr Phe Trp Lys  
                   65                  70                  75  
 Asn Phe Lys Ser Lys Leu Gly Phe Ile Asn Trp Asp Ala Ile Asn  
                   80                  85                  90  
 Lys Asn Gln Val Pro Pro Pro Ser Thr Arg Ala Leu Leu Tyr Phe  
                   95                  100                 105  
 Ser Arg Leu Trp Glu Asp Phe Lys Gln Asn Thr Pro Phe Leu Asn  
                  110                 115                 120  
 Trp Lys Ala Ile Ile Glu Gly Ala Asp Ala Ser Ser Leu Gln Lys  
                  125                 130                 135  
 Arg Ala Gly Arg Ala Asp Gln Pro Gly Ala Gly Trp Gln Glu Val  
                  140                 145                 150  
 Ala Ala Val Thr Ser Lys Asn Tyr Asn Tyr Asn Gln His Ala Tyr  
                  155                 160                 165  
 Pro Thr Ala Tyr Gly Gly Lys Tyr Ser Val Lys Thr Pro Ala Lys  
                  170                 175                 180  
 Gly Gly Val Ser Pro Ser Ser Ser Ala Ser Arg Val Gln Pro Gly  
                  185                 190                 195  
 Leu Leu Gln Trp Val Lys Phe Trp  
                  200

<210> 60  
 <211> 161  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7509086CD1

<400> 60  
 Met Leu Val Ala Ala Ala Ala Glu Arg Asn Lys Asp Pro Ile Leu  
   1                  5                  10                  15  
 His Val Leu Arg Gln Tyr Leu Asp Pro Ala Gln Arg Gly Val Arg  
                   20                  25                  30

Val	Leu	Glu	Val	Ala	Ser	Gly	Ser	Gly	Gln	His	Ala	Ala	His	Phe
				35					40					45
Ala	Arg	Ala	Phe	Pro	Leu	Ala	Glu	Trp	Gln	Pro	Ser	Asp	Val	Asp
				50					55					60
Gln	Arg	Cys	Leu	Asp	Ser	Ile	Ala	Ala	Thr	Thr	Gln	Ala	Gln	Gly
				65					70					75
Leu	Thr	Asn	Val	Lys	Ala	Pro	Leu	His	Leu	Asp	Val	Thr	Trp	Gly
				80					85					90
Trp	Glu	His	Trp	Gly	Gly	Ile	Leu	Pro	Gln	Ser	Leu	Asp	Leu	Leu
				95					100					105
Leu	Cys	Ile	Asn	Met	Ala	His	Val	Ser	Pro	Leu	Arg	Cys	Thr	Glu
				110					115					120
Gly	Leu	Phe	Arg	Ala	Ala	Gly	His	Leu	Leu	Lys	Pro	Arg	Ala	Leu
				125					130					135
Leu	Ile	Thr	Tyr	Gly	Val	Ser	Gly	Pro	Val	His	Gly	Gly	Pro	Pro
				140					145					150
Ala	Pro	Ser	Pro	Ser	Cys	Cys	Leu	Leu	Pro	Cys				
				155					160					

<210> 61  
 <211> 81  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7506914CD1

Met	Val	Gly	Lys	Met	Trp	Pro	Val	Leu	Trp	Thr	Leu	Cys	Ala	Glu
1				5					10					15
His	Leu	Gly	Asn	Lys	Ser	Thr	His	Arg	Arg	Asn	Leu	Arg	Leu	Pro
				20					25					30
His	Asp	Leu	Arg	Ile	His	Ser	Gln	Gly	Asp	Arg	Arg	Cys	His	Leu
				35					40					45
Cys	Gly	Asn	Ser	Ala	Gly	Arg	Ser	Ser	Gly	Phe	Ala	Gly	Lys	Glu
				50					55					60
Cys	His	Pro	Ala	Leu	His	Leu	Pro	His	Phe	His	Leu	Gln	Ser	Arg
				65					70					75
Gly	Thr	Tyr	Ser	Met	Gly									
				80										

<210> 62  
 <211> 214  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 5606114CD1

Met	Gln	Leu	Thr	His	Gln	Leu	Asp	Leu	Phe	Pro	Glu	Cys	Arg	Val
1				5					10					15
Thr	Leu	Leu	Leu	Phe	Lys	Asp	Val	Lys	Asn	Ala	Gly	Asp	Leu	Arg
				20					25					30
Arg	Lys	Ala	Met	Glu	Gly	Thr	Ile	Asp	Gly	Ser	Leu	Ile	Asn	Pro
				35					40					45
Thr	Val	Phe	His	Ser	Cys	Cys	Pro	Gly	Trp	Ser	Ala	Met	Ala	Arg
				50					55					60
Ser	Trp	Leu	Thr	Ala	Thr	Ser	Ala	Ser	Arg	Val	Gln	Ala	Ile	Val
				65					70					75
Leu	Pro	Gln	Pro	Pro	Glu	Leu	Leu	Gly	Leu	Gln	Ile	Val	Asp	Pro

	80		85		90
Phe Gln Ile Leu Val	Ala Ala Asn Lys	Ala Val His Leu Tyr	Lys		
	95		100		105
Leu Gly Lys Met Lys	Thr Arg Thr Leu	Ser Thr Glu Ile Ile	Phe		
	110		115		120
Asn Leu Ser Pro Asn	Asn Asn Ile Ser	Glu Ala Leu Lys Lys	Phe		
	125		130		135
Gly Ile Ser Ala Asn	Asp Thr Ser Ile	Leu Ile Val Tyr Ile	Glu		
	140		145		150
Glu Gly Glu Lys Gln	Ile Asn Gln Glu	Tyr Leu Ile Ser Gln	Val		
	155		160		165
Glu Gly His Gln Val	Ser Leu Lys Asn	Leu Pro Glu Ile Met	Asn		
	170		175		180
Ile Thr Glu Val Lys	Lys Ile Tyr Lys	Leu Ser Ser Gln Glu	Glu		
	185		190		195
Ser Ile Gly Thr Leu	Leu Asp Ala Ile	Ile Cys Arg Met Ser	Thr		
	200		205		210
Lys Asp Val Leu					

<210> 63  
 <211> 36  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7503282CD1

<400> 63
Met Gly His Ala Leu Cys Val Cys Ser Arg Gly Thr Val Ile Ile
1 5 10 15
Asp Asn Lys Arg Tyr Leu Phe Ile Gln Lys Leu Gly Glu Gly Glu
20 25 30
Val Arg Cys Gly Met Arg
35

<210> 64  
 <211> 64  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7503284CD1

<400> 64
Met Gly His Ala Leu Cys Val Cys Ser Arg Gly Thr Val Ile Ile
1 5 10 15
Asp Asn Lys Arg Tyr Leu Phe Ile Gln Lys Leu Gly Glu Gly Glu
20 25 30
Ser Cys Val Thr Ser Ser Arg Thr Gly Arg Arg Pro Ser Glu Lys
35 40 45
Pro Thr Cys Ile Ala Ser Ser Ile Thr Pro Thr Ser Phe Ala Ser
50 55 60
Trp Leu Thr Val

<210> 65  
 <211> 142  
 <212> PRT  
 <213> Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7510501CD1

&lt;400&gt; 65

Met	Gln	Leu	Thr	His	Gln	Leu	Asp	Leu	Phe	Pro	Glu	Cys	Arg	Ile	
1				5					10					15	
Val	Asp	Pro	Phe	Gln	Ile	Leu	Val	Ala	Ala	Asn	Lys	Ala	Val	His	
				20					25					30	
Leu	Tyr	Lys	Leu	Gly	Lys	Met	Lys	Thr	Arg	Thr	Leu	Ser	Thr	Glu	
				35					40					45	
Ile	Ile	Phe	Asn	Leu	Ser	Pro	Asn	Asn	Asn	Ile	Ser	Glu	Ala	Leu	
				50					55					60	
Lys	Lys	Phe	Gly	Ile	Ser	Ala	Asn	Asp	Thr	Ser	Ile	Leu	Ile	Val	
				65					70					75	
Tyr	Ile	Glu	Glu	Gly	Glu	Lys	Gln	Ile	Asn	Gln	Glu	Tyr	Leu	Ile	
				80					85					90	
Ser	Gln	Val	Glu	Gly	His	Gln	Val	Ser	Leu	Lys	Asn	Leu	Pro	Glu	
				95					100					105	
Ile	Met	Asn	Ile	Thr	Glu	Val	Lys	Lys	Ile	Tyr	Lys	Leu	Ser	Ser	
				110					115					120	
Gln	Glu	Glu	Ser	Ile	Gly	Thr	Leu	Leu	Asp	Ala	Ile	Ile	Cys	Arg	
				125					130					135	
Met	Ser	Thr	Lys	Asp	Val	Leu									
				140											

&lt;210&gt; 66

&lt;211&gt; 42

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7500444CD1

&lt;400&gt; 66

Met	Pro	Thr	Thr	Ile	Glu	Arg	Glu	Phe	Glu	Glu	Leu	Asp	Thr	Gln	
1				5					10					15	
Arg	Arg	Trp	Gln	Pro	Leu	Tyr	Leu	Met	Ile	Thr	Val	Val	Leu	Asn	
				20					25					30	
Cys	Lys	Met	Leu	Arg	Met	Ile	Ile	Leu	Met	Pro	Val				
				35					40						

&lt;210&gt; 67

&lt;211&gt; 36

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7510297CD1

&lt;400&gt; 67

Met	Ser	Gly	His	Ser	Pro	Thr	Arg	Gly	Ala	Met	Gln	Val	Ser	Gly	
1				5					10					15	
Ser	Arg	Arg	Pro	His	Leu	Asn	Phe	Asp	Pro	Arg	Pro	Gly	Pro	Ala	
				20					25					30	
Pro	Ser	Gly	Ala	Gln	Ala										
				35											

&lt;210&gt; 68

&lt;211&gt; 569

&lt;212&gt; PRT



&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7640560CD1

&lt;400&gt; 68

Met	Ala	Thr	Glu	Gly	Gly	Gly	Lys	Glu	Met	Asn	Glu	Ile	Lys	Thr	1	5	10	15
Gln	Phe	Thr	Thr	Arg	Glu	Gly	Leu	Tyr	Lys	Leu	Leu	Pro	His	Ser	20	25	30	35
Glu	Tyr	Ser	Arg	Pro	Asn	Arg	Val	Pro	Phe	Asn	Ser	Gln	Gly	Ser	40	45	50	55
Asn	Pro	Val	Arg	Val	Ser	Phe	Val	Asn	Leu	Asn	Asp	Gln	Ser	Gly	60	65	70	75
Asn	Gly	Asp	Arg	Leu	Cys	Phe	Asn	Val	Gly	Arg	Glu	Leu	Tyr	Phe	80	85	90	95
Tyr	Ile	Tyr	Lys	Gly	Val	Arg	Lys	Ala	Ala	Asp	Leu	Ser	Lys	Pro	100	105	110	115
Ile	Asp	Lys	Arg	Ile	Tyr	Lys	Gly	Thr	Gln	Pro	Thr	Cys	His	Asp	120	125	130	135
Phe	Asn	His	Leu	Thr	Ala	Thr	Ala	Glu	Ser	Val	Ser	Leu	Leu	Val	140	145	150	155
Gly	Phe	Ser	Ala	Gly	Gln	Val	Gln	Leu	Ile	Asp	Pro	Ile	Lys	Lys	160	165	170	175
Glu	Thr	Ser	Lys	Leu	Phe	Asn	Glu	Glu	Arg	Leu	Ile	Asp	Lys	Ser	180	185	190	195
Arg	Val	Thr	Cys	Val	Lys	Trp	Val	Pro	Gly	Ser	Glu	Ser	Leu	Phe	200	205	210	215
Leu	Val	Ala	His	Ser	Ser	Gly	Asn	Met	Tyr	Leu	Tyr	Asn	Val	Glu	220	225	230	235
His	Thr	Cys	Gly	Thr	Thr	Ala	Pro	His	Tyr	Gln	Leu	Leu	Lys	Gln	240	245	250	255
Gly	Glu	Ser	Phe	Ala	Val	His	Thr	Cys	Lys	Ser	Lys	Ser	Thr	Arg	260	265	270	275
Asn	Pro	Leu	Leu	Lys	Trp	Thr	Val	Gly	Glu	Gly	Ala	Leu	Asn	Glu	280	285	290	295
Phe	Ala	Phe	Ser	Pro	Asp	Gly	Lys	Phe	Leu	Ala	Cys	Val	Ser	Gln	300	305	310	315
Asp	Gly	Phe	Leu	Arg	Val	Phe	Asn	Phe	Asp	Ser	Val	Glu	Leu	His	320	325	330	335
Gly	Thr	Met	Lys	Ser	Tyr	Phe	Gly	Gly	Leu	Leu	Cys	Val	Cys	Trp	340	345	350	355
Ser	Pro	Asp	Gly	Lys	Tyr	Ile	Val	Thr	Gly	Gly	Glu	Asp	Asp	Leu	360	365	370	375
Val	Thr	Val	Trp	Ser	Phe	Val	Asp	Cys	Arg	Val	Ile	Ala	Arg	Gly	380	385	390	395
His	Gly	His	Lys	Ser	Trp	Val	Ser	Val	Val	Ala	Phe	Asp	Pro	Tyr	400	405	410	415
Thr	Thr	Ser	Val	Glu	Glu	Gly	Asp	Pro	Met	Glu	Phe	Ser	Gly	Ser	410	415	420	425
Asp	Glu	Asp	Phe	Gln	Asp	Leu	Leu	His	Phe	Gly	Arg	Asp	Arg	Ala	420	425	430	435
Asn	Ser	Thr	Gln	Ser	Arg	Leu	Ser	Lys	Arg	Asn	Ser	Thr	Asp	Ser	435	440	445	450
Arg	Pro	Val	Ser	Val	Thr	Tyr	Arg	Phe	Gly	Ser	Val	Gly	Gln	Asp	450	455	460	465
Thr	Gln	Leu	Cys	Leu	Trp	Asp	Leu	Thr	Glu	Asp	Ile	Leu	Phe	Pro	465	470	475	480
His	Gln	Pro	Leu	Ser	Arg	Ala	Arg	Thr	His	Thr	Asn	Val	Met	Asn	480	485	490	495
Ala	Thr	Ser	Pro	Pro	Ala	Gly	Ser	Asn	Gly	Asn	Ser	Val	Thr	Thr	495	500	505	510

```

Pro Gly Asn Ser Val Pro Pro Pro Leu Pro Arg Ser Asn Ser Leu
    425                                430                                435
Pro His Ser Ala Val Ser Asn Ala Gly Ser Lys Ser Ser Val Met
    440                                445                                450
Asp Gly Ala Ile Ala Ser Gly Val Ser Lys Phe Ala Thr Leu Ser
    455                                460                                465
Leu His Asp Arg Lys Glu Arg His His Glu Lys Asp His Lys Arg
    470                                475                                480
Asn His Ser Met Gly His Ile Ser Ser Lys Ser Ser Asp Lys Leu
    485                                490                                495
Asn Leu Val Thr Lys Thr Lys Thr Asp Pro Ala Lys Thr Leu Gly
    500                                505                                510
Thr Pro Leu Cys Pro Arg Met Glu Asp Val Pro Leu Leu Glu Pro
    515                                520                                525
Leu Ile Cys Lys Lys Ile Ala His Glu Arg Leu Thr Val Leu Ile
    530                                535                                540
Phe Leu Glu Asp Cys Ile Val Thr Ala Cys Gln Glu Gly Phe Ile
    545                                550                                555
Cys Thr Trp Gly Arg Pro Gly Lys Val Val Ser Phe Asn Pro
    560                                565

```

<210> 69  
 <211> 433  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7506087CD1

```

<400> 69
Met Arg Ala Val Leu Thr Trp Arg Asp Lys Ala Glu His Cys Ile
  1      5      10
Asn Asp Ile Ala Phe Lys Pro Asp Gly Thr Gln Leu Ile Leu Ala
    20      25      30
Ala Gly Ser Arg Leu Leu Val Tyr Asp Thr Ser Asp Gly Thr Leu
    35      40      45
Leu Gln Pro Leu Lys Gly His Lys Asp Thr Val Tyr Cys Val Ala
    50      55      60
Tyr Ala Lys Asp Gly Leu Trp Ser Pro Glu Gln Lys Ser Val Ser
    65      70      75
Lys His Lys Ser Ser Ser Lys Ile Ile Cys Cys Ser Trp Thr Asn
    80      85      90
Asp Gly Gln Tyr Leu Ala Leu Gly Met Phe Asn Gly Ile Ile Ser
    95     100     105
Ile Arg Asn Lys Asn Gly Glu Glu Lys Val Lys Ile Glu Arg Pro
   110     115     120
Gly Gly Ser Leu Ser Pro Ile Trp Ser Ile Cys Trp Asn Pro Ser
   125     130     135
Arg Glu Glu Arg Asn Asp Ile Leu Ala Val Ala Asp Trp Gly Gln
   140     145     150
Lys Val Ser Phe Tyr Gln Leu Ser Gly Lys Gln Ile Gly Lys Asp
   155     160     165
Arg Ala Leu Asn Phe Asp Pro Cys Cys Ile Ser Tyr Phe Thr Lys
   170     175     180
Gly Glu Tyr Ile Leu Leu Gly Gly Ser Asp Lys Gln Val Ser Leu
   185     190     195
Phe Thr Lys Asp Gly Val Arg Leu Gly Thr Val Gly Glu Gln Asn
   200     205     210
Ser Trp Val Trp Thr Cys Gln Ala Lys Pro Asp Ser Asn Tyr Val
   215     220     225
Val Val Gly Cys Gln Asp Gly Thr Ile Ser Phe Tyr Gln Leu Ile
   230     235     240

```

Phe	Ser	Thr	Val	His	Gly	Leu	Tyr	Lys	Asp	Arg	Tyr	Ala	Tyr	Arg	
				245					250					255	
Asp	Ser	Met	Thr	Asp	Val	Ile	Val	Gln	His	Leu	Ile	Thr	Glu	Gln	
				260					265					270	
Lys	Val	Arg	Ile	Lys	Cys	Lys	Glu	Leu	Val	Lys	Lys	Ile	Ala	Ile	
				275					280					285	
Tyr	Arg	Asn	Arg	Leu	Ala	Ile	Gln	Leu	Pro	Glu	Lys	Ile	Leu	Ile	
				290					295					300	
Tyr	Glu	Leu	Tyr	Ser	Glu	Asp	Leu	Ser	Asp	Met	His	Tyr	Arg	Val	
				305					310					315	
Lys	Glu	Lys	Ile	Ile	Lys	Lys	Phe	Glu	Cys	Asn	Leu	Leu	Val	Val	
				320					325					330	
Cys	Ala	Asn	His	Ile	Ile	Leu	Cys	Gln	Glu	Lys	Arg	Leu	Gln	Cys	
				335					340					345	
Leu	Ser	Phe	Ser	Gly	Val	Lys	Glu	Arg	Glu	Trp	Gln	Met	Glu	Ser	
				350					355					360	
Leu	Ile	Arg	Tyr	Ile	Lys	Val	Ile	Gly	Gly	Pro	Pro	Gly	Arg	Glu	
				365					370					375	
Gly	Leu	Leu	Val	Gly	Leu	Lys	Lys	Met	Tyr	Leu	Leu	Val	Tyr	Ser	
				380					385					390	
Phe	Ile	Leu	Ile	Val	Lys	Asp	Tyr	Phe	Ser	Leu	Ser	Thr	Asp	Val	
				395					400					405	
Leu	Gly	Asn	Leu	Thr	Trp	Lys	His	Val	Cys	Lys	Lys	His	Tyr	Trp	
				410					415					420	
Val	Phe	His	Leu	Phe	Ser	Trp	Tyr	Tyr	Ile	Phe	Val	Gln			
				425					430						

&lt;210&gt; 70

&lt;211&gt; 1879

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1393336CB1

&lt;400&gt; 70

ctgataatgt	gtttcaagag	caaattatgt	aaagagaatt	gaataggcaa	gagagatctt	60
actgaaggag	aactggagtc	tgtatgttgg	aaatgtataa	gccagtacaa	agacacgtgg	120
aaacttattc	ttactacctg	aaaaatctat	ttttaatcat	taagaattgt	catttcagga	180
actcttagtt	tagatggtaa	caaagagact	ttaggagcaa	agttaatatc	cgtgtgtgac	240
tagcttaaaag	acagtgtact	ccacaagcca	gaagattagt	ctccttccct	gtagcctcat	300
ctagcttact	tattcagact	ttctgttcac	aaggggtttg	ttatcttcat	gaaattttaa	360
tatgtgtaag	ttagaaatga	aattagaaaa	agagggttctt	gctgtaaaca	ctgggcagag	420
tagctcagaa	gaagactgtg	gttcttggat	gccctcatgc	ctttgttcca	gtgcaggtgc	480
ctagtggagg	gtgcagtgtc	cttgcaacat	agaaggaccc	atcttgcagg	agctgatggc	540
ttagctcctg	ggaacgggag	gagaggaagg	gcttgtggac	ccaacctctg	ttgaattgac	600
tggctttttg	agagaccact	gaaactcccc	tccccactcc	accccccatc	tgtgctagcg	660
gaagtctatc	tagaaagaaa	aaaggagaaa	gtgtttaaat	acgtgatgga	cgaatgcccc	720
ctgcctctgg	tatacttttt	cttaacagaa	aagggtcattc	taactttgca	taatcttcaa	780
ttataatgta	gaagactttt	ccaccttcct	aaataattaa	aagcaaaacta	atgatctcta	840
aatatttgcc	ttcctgggtc	ttttgaaaat	ccttttgtct	caaataccac	ttcagtgtctg	900
tgtgcatgaa	tagtctgcca	gctgacttgt	gctatggaa	atgtttaccc	ctctgtatac	960
acttaacccta	tctcctgcca	tttacatttt	agctgtttta	gagaactcca	gttaattcag	1020
aaaatgctta	actttttctg	tgttcattgat	tataaatcta	attttgtttt	atgggttttag	1080
gtctcgtcag	cctatctcgg	ccccctctct	ttcatgtcct	gacaaaaaca	agggttaattt	1140
catcccaacc	ggatcagctt	tctgtcctgt	aaaactttcta	ggccccctct	tacctgtctc	1200
tgaccttatg	ctcaagaact	cccctaactc	tggccagagc	tcagctttgg	caactctgac	1260
cgttgagcag	ctctcatccc	gggttttcctt	tacgtctctt	tctgatgaca	ccagcacagc	1320
gggctccatg	gaggcctctg	tccagcagcc	atcccagcag	cagcagctcc	tgcagggaact	1380
gcagggtgag	gaccacatct	ctgctcagaa	ctatgtgatc	atctaaaaaa	ggggggagctg	1440
gcctccaccc	tatgttccat	ggattcggaa	caagatttca	gacatctgca	tgagtgcacaa	1500
actttctgaa	caccaccacc	accaataata	cttatcagca	tcataaagta	tctcttaaac	1560

```

actgatcttg gcaggggacgg aactcctatt cagcagtttt tgtggaaagc agtaatgctt 1620
gcaaaacgtg tgtgtcatte agcattttta gtggagacta tgcattttcat agtatatttg 1680
acagattagt actgtgtcct gtgttttgtt ccagattcct cagtataaat aagctctata 1740
tcaaaaagtt gctgttctaa atagaaaatg tcttgctgtg ttttgctcta tggaaaatac 1800
tgtaattcag gattatgttt actattgata cagggtgtttg tttctaactt ctgttatata 1860
tacaatgcaa aaaaaaaaaa                1879

```

<210> 71  
 <211> 1453  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1431502CB1

```

<400> 71
atgggcatgg aggggtcttct ccagaaccac cttctcattt tctcctgttt acctcacagc 60
caagaatgag ataataaaaac agagaaagaa atagagacgg gaggggtggg ggcagtttcc 120
acggtttctt caagcacagt tcttgcaggg ttgtctgggt tgcacagaga ggctgttggc 180
tcttggacgt ggggtcctcgg gggaccttcg atttctcgcc ggggtccaggg tggccccggc 240
gggaggagga cctcgccccc attcctggga ttcttctctt ccttcccaca ttgtttgttt 300
gttgacttta aattggactc tgccaggaat tcttttctct cttgggttta gtggtgagca 360
ggaggggactt agcatatcac agagcccccg gcagccaagc taccgggtgt cctggcccc 420
ttcagttccc tcecttgctc tcttcgagga ccccccaact caggcaccca caccacgggg 480
ggctgcggga ccccgagcct cactcgtgcc accggccacc cagcgtcccc tgcattcttc 540
ccgggggactc acacaccgcc tcgaggccac tccgccccct cggatggatg tccaagcctc 600
cagccaggggc cccttccctg cttcatctcc aggggttgacg gaggtcgaag gagcacatgt 660
ctgcttgtag ggcacgccag gtcttgatgg aacttgacgt ctagacctag acggtggctc 720
agcctgaacc ctggacggcc tggagcacag ccatgcaact gcttctaaga caaaaccgag 780
gtgggggatta aagctaccct cccaggccgc ctggctggaa actcatcccc ctctgtcccc 840
agcccctgca gcccccatca ctgccgctga cagccgcagg gtgtggctca cagagggaca 900
gcagggaggg agagacggac ggggtgggat cctggggggc tgtggccttc agagcgtctc 960
actgccccag ccttgatggg gttgtctgta gagcggggat tcagctgcct ggtctgtggc 1020
gactgtccgg tggcccaggg ggtggacgtg actctcctgt gcgactttgt ttttctgagc 1080
cgagatggct ccgctgcact ccagcctggg tgacagagcg agactctcaa aaaaaaaaga 1140
aaagaaaaga aaagaaaaga aaaaaagaaa tatcctaaca tggccaggca cagtggctca 1200
tgctgtgaat ccagaaactt tgggatgcca aggcaggcgg atctcgaagt taggagttca 1260
agaccagcct ggccaacatg gtgaaacctg gtctctacta aacatacaca aaattagctg 1320
ggcgtggtgg caccacctg taatcccagc tactcaggag gccgaggcag aagagacgct 1380
tgagcctggg ggagattgtc ccaactgcatt ctaacctggg tgacagagtg agactcccat 1440
tttaaaaaaa aaa                1453

```

<210> 72  
 <211> 2250  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2445220CB1

```

<400> 72
cgaggtcata ggccggggtca cgcggaaggc gggcagcgca ggcgaaagga ggctgtgtcc 60
tcagctcctc cctcggggacg cgctcccat cctgacgggg actgagagcc caggagagct 120
cggggcgtgt ccaaactgcc gcgcgcctt gctccgggtt tgtcttctc cctccgttcc 180
tcttccgct gtatctagca ttccggttcc tggaaagggt accggagctg cgattggggc 240
tgggaccaaa gttgagtcct ggaaagggga gcctggagag cggcgggtgt gggaggcccg 300
gccagctcga tcgcaggctt ccacctggcg gccaatgctt gtttgctact gatgggtgtg 360
agcaagttag atattctata ccggagactt ctctaacaa aactttttat cagaggatgg 420
ggaaggccag aagatctcaa aagactcttt gaattcagaa agatgattgg aaatcgggaa 480
agatgccaga atctggtttc aagcgattat ccagtaacaa ttgataagat tgaagagcaa 540
tcagattgta agatcttaga tggacacttt gtttccccca tggctcacta tgtgcctgat 600

```

```

atcatgccaa ttgaatctgt tattgcaagg ttccaattta ttgtgcctaa agaatggaac 660
agcaaatata gacctgtatg cattcatctt gctggaacag gagatcatca ttactggagg 720
cgacgaacac taatggcccg tccatgatt aaagaagccc gaatggcttc tttattgtta 780
gaaaaccctt attatggctg caggaaaccc aaggaccaag taagggtccag cttaaaaaat 840
gtgtccgacc tttttgtgat gggaggagct cttgttttag aatctgcagc tctcttcgac 900
tggctagaga ggggaaggta cggcccttta ggaatgactg gaatatccat gggaggacac 960
atggcttcc tgcgggtatc caactggcct aagcccagtc cattgattcc atgcctgtct 1020
tgggtccacag catctggggt cttcactacg ggtgtgttga gtaaatcaat taattggagg 1080
gagctggaaa agcaatatta taccagaca gtttatgaag aagaaattat tcacatgctt 1140
gaatactgtg gaacagattc tttcaaaatg ggacaagagt ttgtgaaaca cttcactagc 1200
agtgcagaca agctaactaa ccttaatctg gtttccagaa ctttaaattt agatatatca 1260
aaccaagttg tatcccaaaa acctgctgac tgccataatt ctagcaaac atctgtcagt 1320
gcgacatcag aaggactctt attgcaagat acctctaaga tgaagcgctt caatcaaaca 1380
ctttcaacca acaaaagtgg ttatacaagt cgcaaccctc agtcatacca cctacttagt 1440
aaagaacaaa gcagaaacag tcttcggaaa gagtctttta ttttatgaa aggagtcagt 1500
gatgaatgta ctcattgtag aaatttctca gtcccagttg atccaagcct cattatagtg 1560
gttcaagcca aagaagatgc ctatatcca cgaacaggag ttcgaagttt acaagaaatt 1620
tggcctgggt gtgaaatccg atacttagaa ggggggtcata ttagtgctta tctttttaaa 1680
caaggactct tcagacaagc catctatgat gcatttgacc gcttcctcca taaatacgt 1740
aactagttgg attattatat gtaaccagtg tggccatgat gtttcttaca aaagggagct 1800
tcactctgtg atcatgtgaa ggacaagcag acagcactaa tatatcta atcgctatcaat 1860
ttgggtctgga attcattgtt acacatcagt taactataaa cttcgaatac atttgaataa 1920
tttcaagata atatttgaag taaataatgt taaagtgttt ttactattgt ttattggaat 1980
cccatatatt cagtgttttag tctgttggtta ctgtttatga aaaactaaat ttttaatcat 2040
gaataattta ttaaattaaa taaatttatt catagaaact tttctgggtt ttagtaaaat 2100
tgctaaatca aaccaagatt ttaatatatc agcatttact ttgtaccctt ataacatttc 2160
actcttgact tgatttgaga aacatgttta ttttttgaat gttcttgtgt tcaactaatt 2220
ttctgtctgt gaaataaata cttatgatat 2250

```

<210> 73

<211> 1652

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 5504385CB1

<220>

<221> unsure

<222> (1) ... (1652)

<223> a, t, c, g, or other

<400> 73

```

tataggggaat ttgggtccctc gagcagataa ttcggcagag ggggtgcagaa agaaggcggg 60
cgctttcctg gaagagaagt gggtcaggcc caactgaaaa gcaagcctgg caccagaagg 120
gggacattcc ggaactagat atgaagagac ttttctccct gaagtccccc ggaccgaatt 180
agatgttgta cccaggctgg ttaccctggt aacaggcgctc gcggtttgtc aaggaaactga 240
gagaagagggc ttaagaaaat tattcaccac tcataagtca tggagacatc ctcaaaagggt 300
ctgcttaccc aagttactca gttttggaac ctctagatg atctagctca gagtgacct 360
gagggctatg agaagtttat tcagcagcag ctgaaagaag ggaaacagct ctgtgctgcc 420
ccagaaccac agctttgtct acagaccagg attctgaaac caaaagaaaa aatacttttt 480
atcaacctgt gtcagtggac aaggatccca gctcccaat caaccactca tccagtacct 540
ctaactgttg gcaaaccaga agatacaact gagatatcag atgcttacac agtcattgat 600
ttgacctaca atcctgatgt tcttcattgca gcagaaaagg accaagtga aaaaaatcag 660
ttaattcaga tggccatgaa atgcattgag gagaaattcc agttcaccct ctcacactct 720
taccatatta ccaaatttag aataaaagga agcattcaaa gaatgaaaca aaatctgatg 780
ggaatccaaa ctgattccat agatttaaga gaaaaaatga gaagggaact aactcttgga 840
cagatacgaa gcagtactat gagcaatcca gatcactttc ctcaactgtt actgccaaa 900
gaccaagttt caggcaaagc agtgtgtctg atagaagaga tttccagtac tgaattccag 960
gtggagatga agatgccagc ctatgaacta aaaattgtgc atgatcacag tgaaaaacct 1020
ctgaaaattg agttgaaagt tgaattacct ggtattaatt ctgtctctct ctgtgacct 1080
agtgtttctg aggtaagttg aatagataac actgtaaatt tgtattggcc agcattcact 1140

```

```

aacatatatt atgccacagt aacaaactac ataatttcag gggcttacaa tgatttgctc 1200
aggttacccg gtgaaggagc aattcccttc tgagatgtgc cattcacatc tcagcaaagg 1260
gaaaatagaa tgatcaaacc atgcagtgga tcgtaatact tctactcaga catgacactt 1320
catgatataca cttctactca cgctccactg gccaaagtaa gtcacaaggc cagtctctaca 1380
tcagtaggac tgagaagtgt attctcacag tcacaggaag tcatatggca gggtaaactg 1440
actatcctct tacagaagag caaatagtta gaggcagtc tacagtctac cagtctacac 1500
tcttggttgc aaatatctgc ttctttctgg aaggaaaaca cctccaatcc cagtatcaga 1560
ctccaaagga cagaattntg ctttctcttc ctcttgaacc atggactaag acatgtaaaa 1620
aaagttatgt actccctcag tcccatctgt ca 1652

```

<210> 74

<211> 5821

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 6974948CB1

<400> 74

```

cctctacca ctgcctcttc ctcccgtctt ctacctcctt gcagcgccaa gccggctccc 60
caacctgggtc tgctgaacgc aaaccgtgta gaggtttctg caccgcccc tggggccccc 120
gggtgccgag tcccgttaag gcaaagacgc cagcaagcga ggaagcgag cggaagaaaa 180
acaagcgggc gcgcgagggg agccccagga gggctgccga gtggctggca ggcggctccc 240
gccccctccg ggtggcctcg ccggcggtcg acggcccgga ggacgcgcgc tccgaaaagt 300
tagactccgg gcggcggcgc gctacggcgc gtccctctgc gtcccttctt cgtgcgctc 360
cgggaaagga acttgcttct tcggccccag gattggggag tgtgccgcgg gcccgaggg 420
ggggaacagc cagagggaga cgcgaaccag gccgctggcg gagggctcag ggggagcgca 480
gaggcgacgg ccccatgcc ggcggccaac atgatggaga accggccgct gcccgccctg 540
caggtgcccc agccgcaggg tgcgcccag ggcagcccg tctggtccag tctgtcgacc 600
cccacgcttc gccgcccggc cttcaagatg cgcgcgatga agaacgtaca ggagcagagc 660
ctggaggccg ggctggcccg ggacctgccc gccgtcttgg ccccgggcaa ggagtccctg 720
cagctgccgt ccacgcgat caccgcttcc agcgacgagg acaccccggt gtccaactgc 780
tccacaccca gcgcgtcccc gcgcccga aaa cgcttcttcc tccgcaagt gctgaggggt 840
agggagcggg aggagtgcag tgaaagcagc agccaacaga gcagtacga gagcagccac 900
gatgatgatt ctacaggtt cctgagtcct cgagcgcggg aagaaagcac tgccagtaac 960
tccaaccgca gcacgcgggc ctgctcgccc atctccgga agcggtctcg ctgcgaacc 1020
ccgcagaacc aggcggaga caccatggtg gagaagggtc cagatcactc ctcggaacag 1080
tccccgtcca caccggagca gggcgtgcag cgcagctgct cctcccagtc cggccggagc 1140
ggcggcaaga attccaagaa aagccagagt tgggtataat tgtaagccc cacctacaag 1200
cagagaaatg aagacttcag aaagctcttt aagcagcttc cagacacgga gcgcctcatt 1260
gttgattact catgtgcaact ccaaagagac attctccttc agggccgact ctacctctct 1320
gaaaattgga tctgcttcta cagcaacatc ttccgtggg aaactctgct gacagtccgt 1380
ttgaaagaca tctgttccat gactaaagaa aaaacagctc gcctcattcc caatgccatc 1440
caagtttgca ctgattcaga aaagcacttc ttcacttcgt ttggggcccg ggataggaca 1500
tatatgatga tgttccggct ctggcagaat gctctccttg aaaagcctct gtgtcccaag 1560
gagctctggc actttgttca ccagtgcctt gggaaacgaat tgggcctgac cagtgatgac 1620
gaggactacg tgccccctga cgacgacttc aacacaatgg gatactgtga agagatccct 1680
gtggaagaga atgaagtga tgacagctca tccaagagca gcatagagac caagccagat 1740
gccagtccac agctgcccaa gaaatccatc accaacagca cactaacatc cacagggagc 1800
agtgaggccc ccgtctcggt tgatgggctg cccctggagg aagaggcgct ggagggagac 1860
gggtcccttg aaaaggagct cgccattgac aacatcatgg gggagaagat tgagatgatc 1920
gctcctgtga actcccttc actggacttc aatgacaatg aggacatccc cactgagctc 1980
agtgactctt ccgacacaca cgatgaagga gaggtccagg ccttctatga ggacctgagt 2040
ggcggcagtg acgtgaatga agtcttcaac ttacgcgtgg acaagctcta tgacctctc 2100
ttcaccaact cgcccttcca gcgggatttc atggagcagc ggcgcttctc tgatatcatc 2160
ttccatccat ggaaaaagga ggagaatgga aaccagagcc gaggattctt ttacaccatc 2220
acccttacca accctctggc tcccaaaact gccactgtca gggagacaca gaccatgtac 2280
aaggcgagcc aggagagtga atgttacgtg atagatgccg aagtctcac ccacgacgtg 2340
cctaccatg actacttcta cacaatcaat cgctacacgc tcaccggtgt ggctcggaac 2400
aagagccgac tcagggtctc cacagagctg cgctatcgaa aacagccctg ggggttagtg 2460
aaaacgttca tcgagaagaa cttctggagt gggctggagg actacttccg ccatttagag 2520
agcgagctgg ccaaaacgga gagcacttat ttggctgaga tgcacagaca atctcccaaa 2580

```

```

gagaaggcca gcaagactac aacgggtgcgg aggaggaagc gtcccatgc ccacctgcga 2640
gtccctcacc tgggaagaggt gatgagcccg gtcaccacgc ccacagatga ggatgtgggc 2700
cacaggatca aacatgtggc aggttccaca cagacgcggc atatcccga ggacaccccc 2760
aacgggtttcc acctgcagag cgtgtccaag ctgctgctgg ttatcagctg tgttctggtg 2820
ctgctgggtca tccttaacat gatgctcttc tacaactctt ggatgttggg ataccacacg 2880
cagacctca ctgctggca gggctaaagg ctccaagaaa ggttaccca gtctcagaca 2940
gaatgggccc agctcttaga gtcccaacaa aagtaccacg atactgagct ccaaaaatgg 3000
agggaaatca tcaaactctc agtgatgctc cttgaccaga tgaaggactc gctcatcaac 3060
cttcagaacg gcacaggtc ccgcgactac acgtcggaaa gtgaagaaaa gaggaatcgc 3120
tatcattgac aaggcaggaa cagggtggct gcaagaggcc tgtgcaatac atgtacatag 3180
accatataaa tatatatata taaatatata tatatacaga atataaatat atatatata 3240
tacagatttt aaaaaagaga taatgcctat gtaccaggga gaaggagcgg gccctcccg 3300
gccctgtgct ggccggagca gcgttttctt atgggtggagc agctgaggag ggcaggaaac 3360
gcctctcagc accgacctcc cctgatctcc ctccctccac cctctgttcc ccacctctc 3420
ccttctggc cattcttggc ttttagaagg gaaatgttga gccaaagtta tgcctgcga 3480
gacctaaagg tctcaaaaag aagtcttaag acggcattgc ttaagggtgt tcattcccta 3540
atcccccttt gatattgttc caaaataaaa gagaatcttt tcttccctac ccaggcttc 3600
cccaggcatt ctcttgggaa cacagaagca tcaagggagc ccagtctaaa ctctagggt 3660
tggccgcccc cttgatccag accgtactga gtgtcagatg gagaggatgc ttggaaagct 3720
aggaaggtcg cattgtgtgg tggcaggctc ccgagactcg tgttctgaag gatactctcc 3780
tccatcacc ccacaccttc ctccgacctg tgtagggtgt ggaccagta agggctggga 3840
atttcttact ttccacctcg ctttctctct ggtcttctcc caggctggat ctgggcgaag 3900
tgtcactttt tagtgcttaa agtcctattg tttctctgtg ccctaagagc acattgttta 3960
ttagccaggg tggagcgttt tcgaccttat taaacctctt ttagtgcttc taagcaagtc 4020
aggtctgtgg ctctgaggcc ttacttaga gctgtcagtg tgatttccag gatgagtcag 4080
gggtgtgctgg gccctccagg atttgagaga gggagagaag gcccttgcca cgtccgggg 4140
gcagggtcta cccacagaac actgtgagtg gtgagcacca agatgggttg ccagcctcat 4200
ttgctctttg aacgaaccaa cattagccag tggagacgac tgtctgagac agtttgcaca 4260
tagaaaaagc atctgagcca ctggagggaa ggatttggca aaggtgcatg tgtttgggca 4320
ggaaaaggcc aagtgaatc tatgccctac gggagccatt ccaatcctg agtttctcag 4380
acttggaat ggggggttctg gcgggggttg gagtacaaat cctgctctgg tctctggccc 4440
cacagaggtg taagatcatt aagaagtaga aggcactagc ttcaaagagt gagacattca 4500
gggccttttc agcattggcc cactagcacc atgcatttat ctgaattcac gggccctct 4560
gggcacattc atcatccctc atacactaac tactgtttct tcatgtaatg accaaaagag 4620
tcccaccagc tccaggcaga gcttgtggag attcatgaaa gagtcacagc cagtttggca 4680
tgcagaggtc ttcagtcttg cttattcttc tcctgggctt cttgacacag gggcagaaga 4740
aaccacagat gtcccagatg tcccagaaca tcagcccaaa agacagaaga ttttatcact 4800
catctacccc aaagataaag ttgtattttt ctcttagaat ggcctttaat tccctgtggc 4860
aattacagtt tttgactttt ccatgtgagc aataaggcct tggattttcc ctaagggcag 4920
gatcagaaaa caggaaagtc catgcccagg tgatcctcca agggaggagg gtgaggaggc 4980
tccagggcat gagctagagt ggatattcct cccctagaag aggagatggg aggaaatgca 5040
gtttcatgta atccctgtga agagtgccta tgcccacagc ccagtgggtc cttgttcagt 5100
tcttgtccag gcctttgggt gagggacact ggtagggatc aaggctcatg ggctctgagc 5160
caaagaagga ggcattggtg agactgctgg aaatctcccc cttggcacag ccagaagcca 5220
aaactattcc cagaaaagtt tgaatgcaaa acttgtagat tcctgaattt attccataca 5280
taattatttc cttaagttat ttaattatta aagggtccac ggtgggacct ttgggatgac 5340
ttaaaaaggg ttctttgtgg gagtcaagtg acagagcaga cctggccgca gggagttgtt 5400
ctgcttctgc tttacctctc aggaccattc ccaggaggaa aagggtggcag agagggcagt 5460
actgagtatg tcagaagagc cctcctgggt ttttagcgat ttgaaccac tagacaatgt 5520
gtgcacatag ggaatgtcct ttaagaatca tggatctctt gtaggctctc tgtgttccaa 5580
ataatttttg gactatcata ccaggaaaaa aacaaaatct attcattccc aaaaaaatc 5640
ttccccttag tcatagtggg gaagaaatct tatcaccttg gatagtttgg ggtggaacaa 5700
taaagtgggt cctggcattg ctggtagtgt ggtcaggggc taatcctttt tttatttcag 5760
gcggttaagg gttttggggc cggtttggat ttgtgatggg tggccaaagg tagggaaagt 5820
a

```

&lt;210&gt; 75

&lt;211&gt; 3063

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7501636CB1

&lt;400&gt; 75

```

cggctcgagc ggagctggag ccggagctga agccggagcc gggttggagt ctgggagggg 60
gccggggccgg agcggggctcc agagacatgg ggtcgaccga ctccaagctg aacttccgga 120
aggcgggtgat ccagctcacc accaagacgc agcccggtga agccaccgat gatgcctttt 180
gggaccagtt ctgggcagac acagccacct cgggtgcagga tgtgtttgca ctggtgccgg 240
cagcagagat ccggggccgtg cgggaagagt caccctccaa cttggccacc ctgtgtctaca 300
aggccgttga gaagctggtg cagggagctg agagtggctg ccactcggag aaggagaagc 360
agatcgtcct gaactgcagc cggctgctca cccgcgtgct gccctacatc tttgaggacc 420
ccgactggag gggctttcttc tggccacag tggccggggc agggcgagga gggcagggag 480
aagaggatga tgagcatgcc agggccctgg ccgagtcctt gtccttgcc attgttgacc 540
tgctcttctg cccggacttc acggttcaga gccaccggag gagcactgtg gactcggcag 600
aggacgtcca ctccctggac agctgtgaat acatctggga ggctgggtgt ggcttcgctc 660
actcccccca gcctaactac atccacgata tgaaccggat ggagctgctg aaactgtctg 720
tgacatgctt ctccgagggc atgtacctgc cccagctcc ggaaagtggc agcaccaacc 780
catgggttca gttcttttgt tccacggaga acagacatgc cctgccccctc ttcacctccc 840
tctcaacac cgtgtgtgcc tatgacctg tgggctacgg gatccccctac aaccacctgc 900
tcttctctga ctaccgggaa cccctgggtg aggaggctgc ccagggtgctc attgtcactt 960
tggaccacga cagtgccagc agtgccagcc ccactgtgga cggcaccacc actggcaccg 1020
ccatggatga tgctgatcct ccaggccctg agaacctgtt tgtgaactac ctgtcccga 1080
tccatcgatga ggaggacttc cagttcatcc tcaagggtat agcccggtct ctgtccaacc 1140
ccctgctcca gacctacctg cctaactcca ccaagaagat ccagttccac caggagctgc 1200
tagttctctt ctggaagctc tgcgacttca acaagaaatt cctcttcttc gtgtgaaga 1260
gcagcgacgt cctagacatc cttgtcccca tctcttctt cctcaacgat gcccgggccg 1320
atcagtctcg ggtgggcctg atgcacattg gtgtcttcat cttgtgtctt ctgagcgggg 1380
agcggaaactt cggggtgcgg ctgaacaaac cctactcaat ccgctgtgcc atggacatcc 1440
cagtcttcac agggacccac gccgacctgc tcatttgtgt gttccacaag atcatcacca 1500
gcgggcacca gcggttgag cccctcttcg actgcctgct caccatcgtg gtcaacgtgt 1560
ccccctacct caagagcctg tccatggtga ccgccaacaa gttgtgtcac ctgtggagg 1620
ccttctccac cacctgggtc ctcttctctg ccgcccagaa ccaccacctg gtcttcttcc 1680
tcttgagggt cttcaacaac atcatccagt accagtttga tggcaactcc aacctggtct 1740
acgcatcat ccgcaagcgc agcatcttcc accagctggc caacctgccc acggaccgc 1800
ccaccattca caaggccctg cagcggcgcc ggcggacacc tgagcccttg tctcgcaccg 1860
gctcccagga gggcacctcc atggagggct cccgccccgc tgccccctga gagccaggca 1920
ccctcaagac cagtctggtg gctactccag gcattgacaa gctgaccgag aagtcccagg 1980
tgtcagagga tggcaccttg cggtccttg aacctgagcc ccagcagagc ttggaggatg 2040
gcagcccggc taagggggag cccagccagg catggaggga gcagcggcga ccgtccacct 2100
catcagccag tgggcagtg agcccaacgc cagagtggtt cctctcctgg aagtccaagc 2160
tgccgctgca gaccatcatg aggctgctgc aggtgctggt tccgcagggt gagaagatct 2220
gcatcgacaa gggcctgacg gatgagctct agatcctgcg gttcctgcag catggcacc 2280
tggtggggct gctgcccgtg cccaccccca tctcatccg aatgtggacc cccctgtctg 2340
gtacgacacc gacgtgaagc tgtttgagat acagcgggtg tgaggatgaa gccgacgagg 2400
ggctcagctc aggggaaggc agggccttgg tccctgaggg tcccccatc caccattctg 2460
agctttaaat taccacgatc agggcctgga acaggcagag tggccctgag tgcatgccc 2520
tagagacccc tgtggccagg acaatgtgaa ctggctcaga tccccctca cccctaggct 2580
ggactcacag gagccccatc tctggggcta tgccccacc agagaccact gcccccaaca 2640
ctcggactcc ctctttaaga cctggctcag tgctggcccc tcagtgccc cccactcctg 2700
tgctaccagg cccagaggc agaagccaat gggctcactg gccctaaggg gtttgaccag 2760
ggaaccacgg cgtgtccctt gaggtgcctg gacagggtaa gggggtgctt ccagcctcct 2820
aaccacaaagc cagctgttcc aggctccagg ggaaaaaggt gtggccaggc tgctcctcga 2880
ggaggctggg agctggccga ctgcaaaaagc cagactggcg cactcccgt tcttggggc 2940
atgggtgtgg gtggtgaagg tctcctgcta tattctcctg gaatcgtgga aataacctgg 3000
ctccccttta cccagtaaag aggggaccgg gacggaactg tgaagccacc gatttaatcc 3060
tcc

```

&lt;210&gt; 76

&lt;211&gt; 3100

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature



&lt;223&gt; Incyte ID No: 2535717CB1

&lt;400&gt; 76

```
aggaaatcgg gaaaggagga atagagaggg tccccgagtt cattcaagag cagaccgaga 60
gatgctccga tcccagagtc ggccgggtggc ggggtcaggc cggcgccga ggggggcgaa 120
cgcgagacgc aggtcgcgcg gtttaaactc cccgcctccc gcagcgggag cctgacacct 180
ccacgcactg gtgcggaggg gcaactggct cgggagggcc gggagtcggc gcggggctcg 240
ggggcggttg ccggaacggg tgccgggacg gatcaccgta cggccgaacg cgacccgagg 300
aggcggaaga ggcgcgcgt cactccatgg attcctttgg gcaaccaga ccagaagata 360
atcagtcagt agtcagaaga atgcaaaaga aatactggaa aactaaacag gtctttatca 420
aagcaacagg aaaaaaagag gatgagcact tgggtggcgtc tgatgctgaa ctggatgcta 480
aacttgaggt ttttactct gttcaagaga catgcactga acttctgaag ataatcgaga 540
aataccagct aagactcaat gttatatcag aggaagaaaa tgagctaggg ctctttttaa 600
aatttcaagc agaacgggat gcaactcaag ctggcaaaaat gatggatgcc actggcaagg 660
cactttgttc ttcagccaag caaagattgg cctgtgtgac tccctctgtc cgtctgaagc 720
aagaagtagc aacattcagt caaagggcag tatctgatac cttgatgaca attaatcgga 780
tggagcaggc acgcacagaa tacagaggag ctctactgtg gatgaaagat gtatcccaag 840
agctggaccc agacacctta aagcaaatgg aaaagtttag aaaagtacag atgcaagtga 900
gaaatagcaa agcttctttt gacaagttaa agatggatgt ttgtcagaaa gtggatttac 960
ttggagctag tcgctgcaat atgctatctc attcgtcac tacctaccag aaaacagctc 1020
gaatgatgtc ccaaattcat gaagcctgta ttggctttca tccgtatgat tttgtagctc 1080
tcaagcaact acaagacacg ccaagcaaga ttagtgaaga caataaagat gaacaaatag 1140
gcggttttct tactgaacag ctcaataagc tagttttgtc tgatgaggaa gcaagctttg 1200
agagtgaaca agactgggtc tcccaagagg aatcagaact ttgtctttca cactctgata 1260
accagccagt gccttcacag agtccaaaga aattaacaag atcccccaac aatggcaacc 1320
aagacatgtc agcctgggtc aatctgtttg cagacttgga tccactttca aaccagatg 1380
ctattggaca ctcagatgat gaacttctta atgcttgact gaagttataa tgtcacttca 1440
gtggccttga gacatcaatt ttgcaacgta tttccttcgt ggaaaggatt tagattgtaa 1500
ccgcacaca aaagcacggg gtttgtgaat ataacacctg tcagccaact ttagacagat 1560
ggtaaagacc acatttgaat aagtacacat ctttcatatc ttggatttgc agctgttgg 1620
actatgtgga aaatattaga aacttctatg tggaaaatat tagaaactac agagtttgcg 1680
atatttagat actgaaattt atgtcaaaat aacggctagg aataattctg tcaatatgga 1740
gttgagctta tttcttttga aaccttttta agttgccttg ctggctgtga gaattttata 1800
tgtggataac aaagatagat agatagcatg taaattgggt tgtggtttgg ggtcagtttt 1860
taaatagaat agtagcgagg aggattttct gttttggaaa acaccattag aaccagacca 1920
gctttgtttt ggggttagaga gagtaagatt tgagaactca gtttgcttta atgaaatcac 1980
agagaaactt ggtacttgtt tttcttcatt tggaggctaa aatgtaatgt tttttcattc 2040
atacaaaaata atggacactc cctaattcca ttattaaatc ttgaagggga agtagcagga 2100
taattaattt gctaagccca tccctctgcag aaacagaaaa atctatcttc ccatctccta 2160
aaactcagaa tgcacagtaa tacttaaggc ttgtacaagt gtcttcagac ccactttttc 2220
atacacttgc tatatagtag tatgcagtat ttatattatt cctgaaaaata aaatgagggg 2280
agaatattcc ctaagcaact ggcaatagta ttcttgaaat acctagaaat ttctatctga 2340
atgagggaga cacttatgaa caccttatcc ttacatatat ttgcatactt atctcatatt 2400
ttgtgacata attatttaac ccagaatact ttctggcaga catcacagaaa gctctgtgtg 2460
atcaataagg gagtgtctca tttttctact tccctctttc tgtgggtgac atgatctgag 2520
gttctatttg attactaagc aaaatctgtt acccctacag ggtttagaac ctaagtatta 2580
gagaggaagg ctattttaatg gaagtttagtg taagctgata aaaacgtagc taccgtacac 2640
acacatcaat cactcaattt cctgtccttt taaattgccc accctttaat tttgaagcaa 2700
tttcccaagt gtgtgtttgt tttatatttg taatccagtc cattgcattt ccataagaag 2760
acattttgac tggctgggtg cggtggctca cgcttgtaat ccagcactt tgggaggctg 2820
aggcaggcgg atcacgaggt caggagatgg agaccatcct ggctaacatg gtgaaacccc 2880
gtctctacta aaaatacaaa aaattagcca ggcgtgggtg cacgtgcctg tagtcccagc 2940
tactccggag gctgaggcag gagaatggcg tgaacccggg aggcggagct tgcagtgagc 3000
tgagatggcg cactgcact ccagcctagg caacagagca agactccgtc tcaaaaaaaaa 3060
aaaaaaaaaa aaaaaaagaa agcggagcgg cgcgcgccgg 3100
```

&lt;210&gt; 77

&lt;211&gt; 2210

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 6119548CB1

&lt;400&gt; 77

```

ccggaagtgc ccgcgcaccg gtttccgggtg cagggtgggga aaatggcgggt gtctacagtg 60
ttctcgactt cgtcgctggg aattggtttg tttgttgctg ttggctcagt gtctccatcg 120
gggaaggagg atctgggggt tctactcca ccatcttcac ctcttgctga ggtcaccacc 180
ttgacatcat atcttgataa gttgtccact tctctggatt gggagccttt gggggattag 240
ggatcatgct gtctctgtgt gctcaggatt tgacacagat gcttgctctg tcaaggcaca 300
gcctattgtc tctttgtct agtgtgacat cattcagacg cttctacaga ggtgacagcc 360
caacagattc ccaaaaggac atgattgaaa tccctttgcc tccatggcag gagagaactg 420
atgaatccat agaaacccaa agagcccgcg tgcctatga gagcagaaaag aggggaatgt 480
tggaaaactg cattcttctt agtctttttg ctaaagaaca tctgcagcac atgacagaaa 540
agcagctgaa cctctatgac cgcttgatta acgagcctag taatgactgg gatatttact 600
actggggcac agggagaagg ttttatacaa gaaaatggca cattctgaag tggagctcaa 660
cagacagcaa cctctccag ccctgtggag gaggccggag gttggggccg gagccctggg 720
agcagggcct tgctcgggca gcttctgacc ctcccctcct ggccaggccc cctggggccc 780
tgccccacag catcatgatg gggaaactcc ccctgggggt cgtctcccct tatgtgaaga 840
tgagttcggg gggctacacg gacccctga aattctacgc caccagctac tgcaccgct 900
acggtcggga ggatttcaag ccccggtgtg gcagtcacgt aggcaccggc taaaaatcaa 960
atttccagcc cgtggtctca tgccaagcca gtctggaggc cttagacaac ccggccaggg 1020
gggaacaagc ccaggaccat tccagctctg tggccagcca gagctaccgc ccctggagg 1080
tgcttgacgg caagcatccc ctgccctgga gcatgcgcca gaccagctca ggctatgggc 1140
gggagaagcc cagtgcgggt cccccacca aggaggtccg gaaggtccat ttcgacacc 1200
aggagcacgg gcctcaggcc atcacggggc tggagcccag ggaagtgcct ctgctccacc 1260
agcagcaggg ccaggaccgg ctggagcggg agaacttccg acatggccca cgcttcatga 1320
cgtcggagta caattccaag tatctcaggg accctctaga ccagccagat ttcttgaca 1380
agaaatcgat cggcgcaaaag gaggggagtg gcttcaccaa acagtccac cagagccca 1440
ttgtcttcca accgccctca caggccctcc ctggggaccc tgccctcctc ccaggccaga 1500
gtgtcaccaa gtcagacttc ctcccaaga ctacactaca tggagatgag ttcctacctg 1560
tggtggccag aggtccaaag cgggaaacag ccttcagccg aggggaatgaa aggattctga 1620
acccagagt gccccctccc tgcccagaac ccagcagcgt gagtcaccag cagtttcagc 1680
cacttcacg gatgcaacag acaaacgttg ccctgcttg ccgggagact gtgggaaaaa 1740
aggagccac aggggttcagc cttaacaacc ccatgtatgt ccggagcccc tgtgatcctg 1800
acagggatca gcgatacctg accacctaca accaagggtg ctttgagaac atccccagg 1860
gtctagacca ggaaggctgg actcgaggtg gcatccagcc ccagatgcca ggaggctacg 1920
ccctcagcca gccggtcagc tgcatggagg ccaccccaa ccccatggag agcctgcggc 1980
acctgcatcc ccacgtggga agaaccctga cctcagctga ccccttctac cagaacacac 2040
ctcacagcag ccgctgcgtg gcacacagct gagccctag cccttggttt gtcaatccta 2100
ggggacagga gccaggctac tgtctgtacc cccactccat cccccagggg aacacccctg 2160
cccagctact ttcctaataa aataaagagc attgaagcag aaaaaaaaaa 2210

```

&lt;210&gt; 78

&lt;211&gt; 2460

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 72263451CB1

&lt;400&gt; 78

```

tcgtagctcg gaagtttgct cgagggcgcc cggagaccgc tggggacaga caggtagcgc 60
cgtaaccggc gggatccgca gctccatctg ccgggcgaag gggctctggt cgggagaggg 120
ctcagcgggc ggaagcgctg ctctagagaa ggggtgcggtc gataaccggg ggcccccctg 180
gaagggccgag gatgggcagc agctcgctgt ccgaggacta ccgccagtgc ctggagcgcg 240
agctgcgacg gggccgcgcg ggcgtgtgcg gggacccctc gctgcgcgcg gtgctctggc 300
agatcctggt ggaggacttc gacctgcacg gggcgctgca ggacgacgcg ctggctctgc 360
tcaccgacgg gctgtggggc cgcgcggacc tggcgcccg cgtacgcggc ctggctcgcg 420
ccttcgagct tctggagctc gccgcggtgc acctgtacct gctgccctgg aggaaggagt 480
tcaccacct caagaccttc tctgggggct acgtgcacgt gctgaagggt gtgctctcag 540
acgacctcct cctgaagagc ttccagaaga tgggctacgt acgcagagac agccatcggc 600
tcatggtgac cgccctgccc cccgcctgcc agctggtgca ggtggccctg ggctgcttcg 660
ccctccggct ggagtgtgag atcctgggtg aggtgctggc ccagctgggc accagtgtgc 720

```

tgccagctga	ggagctgctg	caggcacggc	gtgccagcgg	ggacgtggcc	tccgtgtg	780
cctggctgca	gcagcggctg	gcccaggatg	aggagccgcc	acccctgccc	ccccgaggct	840
cccttgctgc	ttacagggcc	ccactggact	tataccggga	cttgcaggaa	gacgaggggt	900
cggaggacgc	cagcctgtat	ggggagccgt	caccaggccc	tgactcgccc	cggcggagc	960
tggcctacag	gccaccactc	tgggagcaga	gtgccaaact	gtggggcact	gggggcccggg	1020
cctgggagcc	cccagctgag	gagctgccgc	aggccagcag	cccaccatat	ggggccttgg	1080
aggaggggct	ggaacctgaa	ccttcgcgct	tctccttcct	ctctctgcgc	cgtgagctga	1140
gtaggcctgg	ggacctggcc	acccctgaaa	gctctgcagc	ggccagcccc	aggcgtattc	1200
gggcagaggg	ggtaccagcc	tcagcctata	ggtctgtctc	ggagccccc	ggctaccagg	1260
cacacagctg	cctgtccctc	ggcgccctgc	ccacctctcg	ctgcgacacc	tgtcgccagc	1320
tgcctgctgc	ccactgtgca	gcccctgccc	cctgcgctcc	aggccactcg	ctgcgtgtgc	1380
tgcttgccga	cgcacagcgg	cgttctgtgg	tacagcgtgc	acagatggac	actctgctct	1440
acaacagccc	cggggcccgg	ccctaggcca	gctggggccc	agggtcaagg	ctcttgggag	1500
acagctcccc	tgggtgcttct	ctttgctggg	gaggggtctg	ggcctctggt	tgccttgtcc	1560
ctgggggagc	cctgggcaca	gcagggggat	ggagagtcca	ggccaaccct	tgctggaggc	1620
gggactctct	gtatatggaa	ggctccaagg	gcagcaggcc	taagccggcc	agcaggagcc	1680
tttccacatt	tccccaccct	cggaaaccag	acgcatgccc	acatgtgctg	ggtagagggg	1740
tccagcctgt	tccagccagt	gcctgaacct	gtccctgaga	aggaccctgt	gccccaaagg	1800
ccgtggctcc	tggcctgtct	cctgcagtag	ggagtacggg	gctcaggagt	tgggggtcct	1860
ccacgtgtcc	cagggattct	ccacacgagt	gtcctcaagg	gcctgttggg	gacaccgagc	1920
tccccccacc	ccaagctctg	ctggggccaa	tctgcccggg	tctgcccggg	gcacggctga	1980
cctggcttct	gggacgcggg	ccctccctct	gggtgtcaca	gcagctgggc	ccacagacag	2040
gaaatgacct	gacctcaaac	cagatgtgga	gggtgcacgag	aggcgcagga	acgggcagtg	2100
gctgacagct	gggaaggccc	tgagcgtcac	tgctgctggg	cgggaaggcc	tgtaggacagc	2160
ccaccagagg	ggcggggctc	tgggtgggtcc	agtggtaaat	gtgccccctc	cctccattcc	2220
ctgtgggtgc	agccacagag	ctggcgagga	gggtgcctct	ccctaccctc	gggtctggcg	2280
agggaggggg	caaggtgcat	ccactgtgac	agccctctcg	ggcccaacaac	tccctgacct	2340
tgggaagctga	gcctgcttct	ctgacctctc	gccccacccc	ctcctcccca	gaacactcag	2400
agggacgggtc	aacatcagag	cacattaaac	gctgggtgaaa	ccaaaaaaaa	aaaaaagggg	2460

&lt;210&gt; 79

&lt;211&gt; 3496

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7502640CB1

&lt;400&gt; 79

taggaagtgg	ctgccgtcaa	tcactcgggg	agactccaaa	cagtgagcct	agagctggag	60
actagcgtaa	accggcgggg	cggccgggtt	tatgaatgaa	gctatggcta	cagattcccc	120
aagaagaccc	agtcgttgta	ctgggtggagt	tgtgggttcgc	cccaggctg	tcacagagca	180
gtcctactag	gaaagtgttg	tgacttttct	gcaggatggt	gtgccacagg	cttacagtgg	240
aacacctcta	acagaagaaa	aggagaaaat	agtctgggtc	agatttgaaa	atgcagattt	300
aaatgatata	tcaagaaatc	tggaaattca	tgaaatacat	agtactggga	gtgaaccgcc	360
tttggttgatt	atgattggct	acagtgatgg	aatgcagggtc	tggagcatcc	ctatcagtgg	420
cgaagcaca	gagctcttct	ctgttcgaca	tggcccaatt	cgagcggcta	gaatccttgc	480
tgtccacag	tttgggtgct	aaaaatgtga	taactttgct	gaaaaaagac	ccctccttgg	540
tgtttgtaag	agcattggat	cttctggcac	aagcccaccg	tactgttgtg	tggatctgta	600
ttcacttcgt	actggggaga	tgggtcaagtc	cattcaattt	aagacaccta	tttatgatct	660
ccattgcaat	aaacggatcc	ttgtcgtagt	cttgcaggag	aaaattgctg	cctttgatag	720
ctgtactttc	acgaagaaat	tctttgttac	aagctgctat	ccatgtccag	ggccaaacat	780
gaatcctatt	gctcttggga	gcccgtgggt	tgcttatgca	gaaaacaagt	tgattcgatg	840
tcactcagtc	cgtgggtggag	cctgtggaga	caacattcag	tcttatactg	ccacagtcac	900
tagtgctgct	aaaacattga	aaagtggcct	gacaatggta	gggaaagtgg	tgactcagct	960
gacaggcaca	ctgccttcag	gtgtgacaga	agatgatgtt	gccatccaca	gtaattcacg	1020
gcggagtcct	ttgggtcccag	gcattcatcac	agttattgac	accgaaaccg	ttggagaggg	1080
ccagggtgctt	gtgagtgagg	attctgacag	tgatggcatt	gtggcccact	tccctgcccc	1140
tgagaagcca	gtgtgctgca	tggcttttaa	tacaagtgga	atgcttctag	tcacaacaga	1200
caccttggc	catgactttc	atgtcttcca	aattctgact	catccttggg	cctctacaca	1260
atgtgctgtc	caccatctgt	atactcttca	caggggagaa	actgaagcca	aagtacagga	1320
catctgcttc	agccatgact	gtcgtcgggt	tgtgggtcagt	actctccggg	gtacttccca	1380

```

cgttttcccc atcaaccctt atggtggcca gccttgtgtt cgtacacata tgtcaccacg 1440
agtagtgaat cgcattgagcc gtttccagaa aagtgtctgga ctggaagaga ttgaacaaga 1500
actgacgtct aagcaaggag gtgctgttag ccctgttcca ggtctatcaa gcagcccttc 1560
tgggtcacc cctgcatggga aactgaacag ccaagactcc tataacaatt ttaccaacaa 1620
caaccctggc aaccctcggc tctctctct tcccagcttg atggtagtga tgctctctgc 1680
acaaatcaag cagccaatga cattggggac catcaccaaa cgaaccggca aagttaaacc 1740
tctccacaa atttcaccca gcaaatcgat gggcggagaa ttttgtgtgg ctgctatctt 1800
cggaacatcc aggtcatggt ttgcaataaa tgcaggctcg aaaagagaaa aagatcagtc 1860
caaacaagtt gtagttgagt ccctgtacat tatcagttgc tatggcacct tagtggaaac 1920
catgatggag ccgcgacccc tcagcactgc acccaagatt agtgacgaca caccactgga 1980
aatgatgaca tcgcctcgag ccagctggac tctggttaga acccctcaat ggaatgaatt 2040
gcagccaccg ttaaatgcaa accaccctct gctcctcgct gcagatgcag tacagtatta 2100
tcagttcctg cttgctggcc tgggtccccc tggaaagtcct gggcccatta ctgcacatgg 2160
gtcttacgac agtttagctt ctgaccatag tggacaggaa gatgaagaat ggctttccca 2220
ggttgaatgt gtaacacaca ctggacccca tagacgtctg tggatgggtc cacagttcca 2280
gttcaaaaac atccatccct caggccaaac cacagttatc tcatccagtt catctgtgtt 2340
gcagttctcat ggtccgagtg acacgccaca gcctcttttg gatthttgata cagatgatct 2400
tgatctcaac agtctcagga tccagccagt ccgctctgac ccgctcagca tgccagggtc 2460
atcccgtcca gtctctgata gaaggggagt ttccacagtg attgatgctg cctcagggtac 2520
ctttgacagg agcgtgaccc tgctggaggt gtgcgggagc tggcctgagg gcttcgggct 2580
gcggcacatg tctccatgg agcacacgga ggagggcctc cgggagcgac ttgccgacgc 2640
catggccgag tcacctagcc gggacgtcgt gggatccgga acagaacttc agcgagaggg 2700
aagcatcgag actctgagta acagctcagg ctccaccagc ggcagcatac caagaaactt 2760
tgatggctac cgatctccgc tgcccaccaa tgagagccag cccctcagcc tcttcccgc 2820
tggcttccc taggtaccag caacctgctt ctgactggcc agccccctcc cctgctggag 2880
gaggggagaa gccccgctct ggtcctaccc ttcagttctt gctcttctt catcaaccac 2940
cttccccaa cttagtgaac gcagccgccc atcctacctg gatggagaag agacccttct 3000
ccaagcacct cagcgcactt gccctctgcc acacctgtcg gtggaggctg tggccaggag 3060
agactgtaga agctcggtcc ctgtgtatgt ttgcatatga catcctgcat tggatccgct 3120
tttgtattht ttaaccatac ccacgggtgg gcyggtgggg ggagccatga tgcttctgca 3180
gcatgccagt ttgttactg agaccttcac agaaggtcgc acaaccaccc tggacgggtga 3240
gaaactaagg tgactgaggc aagtcattga accacctga gccttggctt cctcatctat 3300
gaaatggagt tgacagcatc tgtcctgcct gtggtaaaaa ttgctgtgat cccacaacca 3360
aaaggaaaga gagaattatg agaaaaatat gtacttatct tgtgaaaata tgtgagattt 3420
tcatgagaga cttcctgttc cagtggactc ttccaccgcc cttgagtggg gaacatccag 3480
gccttcacac tcacgc 3496

```

<210> 80

<211> 1334

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7505807CB1

<400> 80

```

ttggtgagga gagtgcaggg aggtgacgcg cgctgccggg gcgagggttc gaggggagg 60
ggtgaagaat gtgtgggcga acatcctgtc acttacctag agatgttctc acgagagctt 120
gcgctacca ggatcggcgg ggccagcagc ggctcccgga gtggagggac cctgataagt 180
actgcccctc ttacaacaag agtccctcaat ccaacagccc agtgcttctg tctcgactgc 240
actttgagaa gtcaggtagc attggtgctg cagatagtcc tgagaactgg gagaaagtct 300
gggacaactg gaggtgctg acaatggccg ggtctttga ctgctgggag ccccagagg 360
gaggagatgt cctgtattcc tataccatca tcacagtggg ttcttgcaaa ggcttgagt 420
acatccacca caggatgcct gccatattag attgagagga ggcagtttct aaatggctt 480
actttggtga agtctcaact cagggaagctc tgaattaat ccaccaaca gagaacatca 540
ccttccatgc agtctcttct gtggtgaaca actcgcgaaa caacactcct gagtgtctgg 600
ctcctgtcga cttggtggtc aaaaaggagc tcagggcaag tggcagtagc cagaggatgt 660
tgacgtggtt ggccacaaaag tcacccaaaa aggaagactc aaaaacacct caaaaggga 720
agtcagatgt tcccagtggt tccagtcagt tcctgcagaa gactccactc cccaccaaga 780
gaggcactgc aggactccta gagcaatggc gaaggaggag gaacctgtgg 840
ccaagcgtcc ttacagccag tgacacagga ctttcagaga ccaaggccag ggtctgctgc 900
actgctgttc tgataatagg ttcttaacat tgtatgtata tgtgtttgct ttgggaggag 960

```

```

gtggcactgt gttagttgac agttgtgggc tcatgtagtc ttttttgcca tgagtaggag 1020
cccctagtggt ggctgggtgga cagcttttggga agaggtgtcc tgctgctgtt accagccatg 1080
tgggccccat agggggcactg cgccctgctgc ccttttcctgg cagggctggt ggagtcttcc 1140
ctcaaagcat gccttaccca gctgggaagt ctctgccctg atctgggtact ccttgtagta 1200
agctgttttc tgctcagcca ctgggctctt tcactttttt agttcttaaa aattttatttt 1260
taagttctaa aataaaataa aaataagttc ttaaaattta tttttttcct gaataaattg 1320
tatttggttaa aaaa 1334

```

```

<210> 81
<211> 2100
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7506413CB1

```

```

<400> 81
ccggaagtgc ccgcgcaccg gtttccgggtg caggtgggga aaatggcggt gtctacagtg 60
ttctcgactt cgtcgctggg aattgggtttg tttgttgctg ttggctcagt gtctccatcg 120
gggaaggagg atctgggggtt tcctactcca ccattctcac ctcttgctga ggtcaccacc 180
ttgacatcat atcttgataa gttgtccact tctctggatt gggagccttt gggggattag 240
ggatcatgct gtctctgtgt gctcaggatt tgacacagat gcttgctctg tcaaggcaca 300
gcctattgtc tcctttgtct agtgtgacat cattcagacg ctctacaga ggtgacagcc 360
caacagattc ccaaaaggac atgattgaaa tccctttgcc tccatggcag gagagaactg 420
atgaatccat agaaacccaaa agagcccgcg tgctctatga gagcagaaaag aggggaatgt 480
tggaaaactg cattcttctt agtctttttg ctaaaagaaca tctgcagcac atgacagaaa 540
agcagctgaa cctctatgac cgcttgatta acgagcctag taatgactgg gatatttact 600
actggggccac aggggagaagg ttttatacaa gaaaatggca cattctgaag tggagctcaa 660
cagacagcaa ctcttcccag ccctgtggag gagggccggag gttggggccg gagccctgga 720
agcagggcct tgctcgggca gcttctgacc ctcccctcct ggccaggccc cctggggccc 780
tgccccacag catcatgatg gggaaactcc ccctgggggt cgtctcccct tatgtgaaga 840
tgagttcggg gggctacacg gacccccctga aattctacgc caccagctac tgaccgcct 900
acggtcggga ggatttcaag ccccggtgtg gcagtcacgt aggcaccggc tacaaatcaa 960
atttccagcc cgtgggtctca tgccaagcca gtctggaggc ctagacaac ccggccaggg 1020
gggaacaagc ccaggaccat ttccagtctg tggccagcca gagctaccgc cccctggagg 1080
tgcttgacgg caagcatccc ctgcccctgga gcatgcgcca gaccagctca ggctatggg 1140
gggagaagcc cagtgcgggt ccccccacca aggaggtccg gaaggtccat ttcgacacc 1200
aggagcacgg gcctcaggcc atcacggggc tggagcccag ggaagtgcct ctgctccacc 1260
agcagcaggg ccaggaccgc ctggagcggg agaacttccg acatggcca cgcttcatga 1320
cgtcggagta caattccaag tatctcaggg accctctaga ccagccagat ttcttgaga 1380
agaaatcgat cggcgcaaaag gaggggagtg gcttcaccaa acagtcccac cagagcccca 1440
ttgtcttcca accgcctcca caggccctcc ctggggaccc tgccctcctc ccaggccaga 1500
gtgtcaccaa gtcagacttc ctcccgaaga ctacacctca tggagatgag ttctacctg 1560
tgttggccag aggtccaag cgggaaacag ccttcagccg agggaatgaa aggattctga 1620
accccagagt gcccctccc tgcccagaac ccagcagcgt gagtccacc cagtttccag 1680
cacttcatcg gatgcaacag acaaacgttg ccctgcttgg ccgggagact gtggggaaaa 1740
aggagccac agggttcagc cttaacaacc ccatgtatgt ccggagcccc tgtatcctg 1800
acagggatca gcgatacctg accacctaca accaagggaac cctgacctca gctgacctc 1860
tctaccagaa cacacctcac agcagccgct gcgtggcaca cagctgagcc ctgagccctt 1920
ggtttgtcaa tcctagggga caggagccag gctactgtct gtacccccac tccatcccc 1980
aggggaacac ccctgcccag ctactttcct aatgaaataa agagcattga agcagaaaaa 2040
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaagatcttt 2100

```

```

<210> 82
<211> 1301
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1283631CB1

```

&lt;400&gt; 82

```

gttaacaaac aggtaaaaat agtgctctgc catccagtat gagaagataa ttttatgaac 60
tatgtcttaa aaatgtctga ggtaatttgg agtgtttgag tatttaatat gttaaagtga 120
aaccaataga tttcctgaga ttccctccagt atttattagg aaatggaaga tggaaaagctt 180
ttcaactaaa aagatgtttt tgatatactt ttcttagatg ggaaaactga caaccatgcc 240
tgcagggtctg atatatgcat ctgtaagtgt acatgcagcc aaacaagagg aatccaaaaa 300
gcagctagtg aaaccagagc agctccccat atatactgca ccaccgctcc agtctaaata 360
tgttgaagag cagcctggtc atttacaat gggttttgct tccatccgca ctgcaactgg 420
ttgttacatt ggctggtgca aggggtgtta tgtctttgtg aaaaatggga taatggatac 480
agtacaattt ggaaaagatg cttatgtcta tctgaagaat cctcctcgag attttcttcc 540
gaaaatggga gttattacag ttccaggatt ggcggttggt gtttcagcga gaaaagggtt 600
caagtttaag aaaattactt atcctctggg actggccact ttaggagcaa ctggttgcta 660
cccagttcag tccgtaataa ttgctaagggt aacagcaaaa aaggtatatg ctacaagcca 720
gcaaattttt ggagcagtta aatcattgtg gacaaaaagc agcaaagaag agtcactccc 780
taaacctaaa gaaaaaacta agctaggatc ctcttccgaa atagaagtac ctgcaaaaaa 840
aactcacgtc ttgaaacact cagtgccttt gccaacagaa ctcagctctg aagcaaagac 900
caaatacaga tccacttcag gtgcaaccca gtttatgcct gacccaagc tcatggatca 960
cgggcagtc caccagaaag atatatagat gtacagcact agaagctgaa gtagactgca 1020
tgcagagact acacagaaaa ctacaagatg tgtggcggtt caaataatga tgaaaaataa 1080
tcagtataat ggtaactgat acatagatga caagtcttct ccttacattt 1140
ccaaatatgc catttgataa attacataag tggttaaac acaaaactgtg aatgcagtg 1200
aaacaatatt aaatgattga tgaggagact taggtagaag tattagcttg catttgatga 1260
cctttaaggg cacttaaaaa aaaaaaaaaa aaaagatctt t 1301

```

&lt;210&gt; 83

&lt;211&gt; 7783

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1740413CB1

&lt;400&gt; 83

```

cgccgcgttg ggcctaactc gagtcctgcc gcctcccggg agtgccgtgc gccgcagccc 60
gggcccaggc cccggcagcg cctgggacaa ggtcttcaga actactagca gataatttgg 120
ggggacttca tattcagaag tctatataaa ggattcctca tgtccataac atgttggatg 180
aggctctgca gctcaccccc actctcagag tggctcagct ccattaattg gaccccgta 240
tttccaatct ctgctgtgtt ggacgtcatg agcattgcaa tccctctggg agtcaccaca 300
tcagatacat cctactcaga tatggctgct ggatcagagg tatcctgtct tggtagtgaa 360
tggaagaggg cctgggtccct ggggcagtg cagaaggcact ttgatcatct cactcactgc 420
aaatgtcatg tgaggtttca tggaagtatt aagagatcca attaagaaga atagttctga 480
gtctaaacca gccagagtg gcttctctag gggaaactcc ccgctcagct gccctgaatc 540
tgtggaggct agtccagcag ttaatgagaa gagcgtgtat tccactcata attatgggac 600
cactcagagg catgggtgtc gaggactgcc ttatgctacg atcatccctc gttctgacct 660
gaatggcctg ccgtcgcctg tagaggaacg ctgtggagac agcccgaact ctgaaggaga 720
aactgtacct acctggtgtc cttgtggtct ttctcaggat ggcttccttc tcaactgtga 780
caagtgcagg ggaatgagca gggggaagggt tattagactt catcggcgga agcaggacaa 840
catatcaggt ggggatagca gtgcaacaga aagctgggat gaggagcttt ctcttccac 900
tgtgttgtat acagcaacac agcacacacc tacaagcatc accttaactg ttagaagaac 960
caaaccacag aagcggaaaa agagtccaga aaagggtcgt gcagcaccaa agacgaagaa 1020
aatcaagaat tctccctctg aagcacagaa tttagatgag aatacaactg agggctggga 1080
aaatcggata agactatgga ctgaccagta tgaagaagct ttcactaatc agtacagtgc 1140
agatgtacag aacgcgcttg aacaacacct acattctagc aaggaatttg tgggcaaacc 1200
tactatttta gacactatta ataagactga attggcctgt aataacacag ttattgggtc 1260
ccaaatgcag ttacagctgg gaagagtcac tcgtgttcaa aagcaccgga agatcctgag 1320
ggctgcaaga gatttggctt tggacactct tataatagag tatcgtggga aagtcagtgt 1380
acgacagcaa tttgagggtc atgggcattt cttcaaaaaa ccatacccct ttgtgctctt 1440
ctactcaaaa ttcaatggtg tagagatgtg tgtggatgcc cgtactttcg gtaatgatgc 1500
tcggttcatc agaagatcat gtacaccaa tgcagagggt cgacacatga ttgcagatgg 1560
gatgattcac ctgtgcatct atgctgtgtc tgccatcacc aaggatgctg aggtcaccat 1620
agcatttgat tatgagtata gtaactgtaa ttataaagtg gactgtgcct gtcacaaggg 1680
aaaccggaat tgtcctatac aaaaaaggaa tcctaagtct acagaactgc cactcctacc 1740

```

acctcctcca	agcctaccca	ccattggagc	agagactaga	cgtagaaaag	cacgacggaa	1800
agagctagag	atgggagcagc	agaatgaggc	ttcagaggag	aataatgacc	agcaatcaca	1860
agaagttcca	gaaaaagtaa	ctgtatccag	tgatcatgag	gaagtagaca	atccagaaga	1920
aaaaccagaa	gaagagaaaag	aagaggttat	agatgaccag	gagaacctag	ctcatagcag	1980
gaggaccagg	gaagatagaa	aggtagaagc	catcatgcat	gcttttgaaa	acttagagaa	2040
aagaaagaag	cggcgggatc	agcccttggg	acagagcaac	tctgatgtag	agattactac	2100
aaccacctca	gagactcctg	ttgggtgaaga	gacaaaaact	gaagcccctg	aatctgaagt	2160
tagcaactct	gtttcaaaatg	ttaccatccc	aagcacccca	cagagtgttg	gtgtgaatac	2220
cgggaggtct	tcccagcgag	gggatattgc	tgcagaaaaa	ctagtcccca	agccacctcc	2280
agcaaagcct	tctaggcccc	ggccgaagag	tcgaatttct	cggtagcagg	ccagttcagc	2340
ccaaagacta	aagcgtcaga	agcaggccaa	tgcacagcag	gcagaattgt	cacaagctgc	2400
cttggaagag	ggagggaagta	acagttagt	aactcctact	gaagctggaa	gtctagacag	2460
ttcaggagaa	aacaggccat	taacagggtc	tgacccaact	gtgggtgtcaa	ttactggatc	2520
ccatgtcaac	cgtgctgcat	ctaaataccc	caaaaccaa	aagtatctag	ttacagaatg	2580
gttgaatgac	aaagcagaga	agcaagagt	ccctgttgag	tgccctttac	gtatcacaac	2640
ggatccaact	gtactggcaa	cgaccctaaa	catgtttacca	ggtcttatcc	attccccgtt	2700
aatttgcacc	acccccaaac	actacattcg	ctttggctca	ccctttatcc	ctgagagacg	2760
tcgaaggccc	cttctgcctg	atggcacatt	cagctcctgt	aagaagcgtc	ggataaaaaca	2820
agccttagaa	gaagggatga	ctcaaacatc	atctgtaccc	caagagacta	gaactcagca	2880
cctatacca	agcaatgaga	atagtagctc	ttctagtatc	tgcaaagaca	atgcagactt	2940
gttagagcca	ttaaagaaat	ggaagtctcg	ctatctgatg	gagcagaatg	tcaccaagtt	3000
acttcggcct	ctgtctccag	tcacaccacc	ccctcccaat	tcagggtcaa	agagtcccca	3060
gctggccaca	cctggctcat	ctcaccagg	agaagaggag	tgctgaaatg	gatacagcct	3120
catgttttca	ccagtcacat	ctcttactac	tgctagtgcg	tgcaaacactc	ctctacagtt	3180
tgagctttgt	caccgaaaag	acctggattt	ggcaaaagta	ggataccttg	actccaacac	3240
taacagctgt	gctgatagac	cttccctact	caactcaggt	cattctgacc	tggctcctca	3300
tcctccctc	ggacccactt	ctgagactgg	tttcccaagc	agaagtggag	atggacatca	3360
gaccctcgtg	agaaactcag	accaggcatt	tcggacagag	ttcaacttga	tgtagccta	3420
ctcccccttg	aatgctatgc	ctcgagcaga	tggactgtat	cgaggatctc	ctctagtggg	3480
ggataggaag	cctttacatt	tggatggggg	atattgttcc	cctgcagaag	gattttccag	3540
catatatgaa	catggcttaa	tgaagacct	ctctcgtgga	tccttgtcac	ctggtgggtga	3600
aagggcctgt	gaaggagtcc	catctgcccc	ccagaaccca	ccacagagga	aaaaagtatc	3660
cctgctggag	taccgaaaac	ggaaacaaga	agctaaggaa	aattctgctg	gtggggggag	3720
tgactctgca	cagagcaaaa	gcaagtctgc	aggagctggg	caaggcagca	gtaactccgt	3780
ttccgacact	ggtgcccattg	gtgtgcaggg	atcctcagcc	cgaactccat	cttccccctca	3840
caaaaaatc	tccccatctc	attcctctat	gtcccatttg	gaggcggtaa	gccccatcaga	3900
ttccagaggg	acttcttcat	ctcactgcag	acctcaagag	aatatcagca	gtaggtggat	3960
ggttcccaca	ctagtgaac	gactccgaga	aggagggagc	atccccaaag	tcctccgaag	4020
cagcgtgagg	gtggcccaaa	agggagagcc	ctctcccaca	tgggagagta	acatcacaga	4080
gaaagactca	gaccctgcag	atggagaagg	cccagagaca	ttaagctcag	cactctctaa	4140
aggagcaaca	gtttacagcc	cttccagata	cagctaccag	ctcctgcagt	gtgatagtcc	4200
tcggacagaa	tcacaaaagcc	tccttcagca	gagttcctcc	cccttcagag	gacatcctac	4260
acagtctcca	ggatacagtt	atcgaactac	gtcactgaga	cctggaaaacc	ccccctctca	4320
cggttcttca	gaatcatccc	tctcttcac	tgcctattcc	agccccgccc	accctgtgtc	4380
cacagactcg	ttggccccat	ttacgggggac	accaggggtat	tttagcagcc	agccacattc	4440
tggaaacagc	actggcagca	atcttccaag	gaggagctgc	ccttctagt	ctgctagccc	4500
taccctgcag	ggaccctcag	actcgccaac	ctcagattca	gtttctcagt	ccagcacagg	4560
aactctgagt	tccacctcct	ttcctcagaa	ctctaggtcg	tcattgccat	cagacttacg	4620
gactatcagt	ctgcccagtg	ctgggcagtc	agctgtctac	caggcctcca	gggtatctgc	4680
ggtttccaat	tcacagcact	acccacaccg	tgggagtggg	ggtgtgcacc	agtaccgact	4740
ccagccactg	caagggtcag	gagtcaagac	tcagacggga	ctttcctagg	gcttctggat	4800
ttgggcaaac	agaactgaat	gagcccatag	ctgcttcctt	ccagctgcct	ctggaacctca	4860
ggccgagcat	attgctgagg	aacggggggt	acaaggtgcc	agaggattgg	gtctggtgga	4920
caagaaacaa	gacttgtggt	cacaattggc	ctctggcctt	ggagaaaagct	gtaaatcttg	4980
tctgaagcag	agactataaa	gaagtttctc	cctgtgtca	aggggtacatt	gttgacaagc	5040
aaatggtgtt	tcggtttagta	acggttctaa	gtgcaatgag	ttgtgttgaa	gcctccgtct	5100
cccaccttg	cctgtagccc	gtagtcactt	gtgcagtgcg	gacatctttt	taaatttaaa	5160
aaaaaaaaaa	aaaaaagttt	tcaaaaggaaa	aaaagttaaa	agagccaatc	tcaaagcccc	5220
aagccactcg	agtactgtta	gggttttatg	cacttaagaa	aaaaggtagg	tatgtaaatg	5280
ttcatcctaa	gacaacatt	ccaaaagcag	gtatctggcc	aatgtgtgtc	caccaagaat	5340
actgtttatc	tttgtcttaa	gatcaccaag	aaataggcaa	ggatagtaaa	gcttgggaacc	5400
tgcaccaact	ggagggtgcc	tggctctttg	aagaaaagct	catggtcagc	tcttgattat	5460
tcgggagcag	attatttgag	tagattgtct	gagcctccaa	ctgttaccat	cctactcccc	5520

```

cttcccaagc tatlctacag ctacagtaacc catgaagtaa gtagacaaga aaaggaggaa 5580
tgagacatga tataggccaa ttgcattgct acttaccagc ttttgccaat aaatttcatt 5640
aggaaggata ccgagtggtt tgggaatgct tcgaatttta ttttttctac tcccaattaa 5700
tcaggagttg atgatcccat gagcaggacc gcctccatga ttggggagca tgcacttggtg 5760
actgcagggt aagagtggtg agatagggtt gtggagtggc accgacagga ctgtgattgt 5820
gtgtgggcct gccccacatt tctctggggg atgcttatgt gagagtgggc ccagtgaag 5880
agttaccaag ccacccacac ccctaacact gttctggatg agagatgaga gcagaccggc 5940
ttctcccat cagtgcattg tgctgttgt acacccctgg aggagccctg gagccagccc 6000
aggtggggta cacaatcttt ttaaattcca tatggttgcc agcttatttc tttcacttgt 6060
ttactgtaat atctggcgtg tttttattta tctaattttg tattcagtta taaccatggt 6120
aggggtagtg aatatatgac aggtgtaatc cctggtgctg cagtggacct tcttttcttt 6180
tggaacaagt aatactgtga gtttccctcc ttcttccct ctaatttgtt ttcttttttt 6240
ccccagcctc ttgcatcccc ttcttttcta cctgtccta caactatcat atgcacagtc 6300
ttctctcttt gtgtgtgact gttacaaaat ttcaactttc aaaatcgaaa tcagggtgtt 6360
gtcctaaatga ggggagattt tttttttttt ttttttttaa atgctgagac ctacgacag 6420
tacttttctt ttgtgtgttt cccccacaaa cccatcagtc tgggagagca ttgggagtg 6480
aaatcatgtt gcctgggatg ctggtttctt tgtatattat ataaaacgta tgtaatgtc 6540
tctccatttg ggctgggggt tgcattctcc ccttggctat ttaaccaagg ggagaggcca 6600
gcgggcaggc ggcctcacc ctgcctggca cgtgcagaga cccagccac tctgtgtggg 6660
cagggtgctg tcaagaccag acctcttggg ggggtagggg cgggggggtg gggggaactc 6720
ttggaaggga agaagtatca cttctttctc aagtggagtg tttacacctt gctgtaacat 6780
ttgaactttc acaagagatg taataatttt gataataaaa ttcttaacca taataatcat 6840
aaagttgaga acttctttgt ccaattttgt ggccttgcct tgctgtgttt tctaatgcc 6900
atctttatct gaggagaatg tgcccttggg ggagtgtgtt tgcccatgat aatggggcca 6960
ggaatactgt ctgatcttcc ttgccttttt atgattgcag tcaggacagt gtgaaggcaa 7020
acctgcagca gggcaagtga acattaaaaa ctgccttgaa cttgagtgtt gttcctgacc 7080
ttggcctcat tcattcagca actatttatt ggtctttttc attcaaaaac aatgctggac 7140
attgtttttg gtgctgggct tatggtgatt agaatgctga ttgggggtgg gcaaagtaag 7200
gtctcttctt tcatggagt tcatcttag tggaggagag aactaaacta gtatataatt 7260
ttaaatagtc attgccatgt taattgtttc atagcagata tgtgactaag acagttgaag 7320
gagccactct agattgggtg gtcaagaaaa gtctctgagg agttgacatt taaactgaga 7380
cttcagtgc aaggaggcag tctgggacg atgtgagtga aattcattct tgaacgttgc 7440
agactctgtg gttagaaaaa cttgggttaag atgtaataaa tggatggggc tgggaaagag 7500
gatagaagat cctgtaagag ggagataggg gacagatcat gtagagcagg agtcggcaaa 7560
ctctttctat aaagggccag atagtaaata tttagacttt gcaagccatg tatggtctct 7620
gtcataacta ttcagccctt tggctgggtg cgaatcatgc ctgtaattcc agcactttgg 7680
gaggcccaag agagcagatc acttgagggtc agaagttcga gaccagcctg gccaacatgg 7740
caaaacctg tctactaaaa atacaaaaat taaccagggtg tgg 7783

```

<210> 84

<211> 1977

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 1951731CB1

<400> 84

```

tgctgctgca cctcttcagt gatgccaaact acaacctgct gggctttaac gcctcattcc 60
gcttctccct gtgcccgggt ggctgccaga gccacgggca gtgccagcca ccgggtgtgt 120
gtgctgcga gccgggctgg gggggtcctg actgtggcct gcaggagtgc tcagcctact 180
tgggcagcca cggcaacctgc gcctcgcccc tgggaccatg ccgctgtgag cctggcttct 240
tgggacgtgc ctgtgacctg cacctgtggg agaaccaggg ggctgggtgg tggcacaacg 300
tgagtccag ggacctgcc ttctctgccc gtattggggc agctggcgcc ttctgtccc 360
caccagggtc gctggcagtt ttcgagggcc aggacctcaa caatgcctg ggtgacctcg 420
tctatacaa cttctccgcc aacacctggg agtcttggga cctgagtcct gccccgcacc 480
agcgggggct attgggagca ggtgattccg gcaggcggac ggccccctgc tgccactggc 540
cactccatgg tgttccatgc cccctcccgt gccctgctgg tccatggtgg acaccggccc 600
tccactgcc gggtctctgt gcgagtgaac tccactgagc ttttccacgt ggatcggcac 660
gtgtggacga cggtgaaggg gcgggatggg cttcagggcc caaggagcg agccttccac 720
acagccagtg ttctgggcaa ttacatggtg gtctatgggg gcaatgtgca caccattac 780
caggaggaaa agtgctacga agatggcatc ttcttctacc acctggctg ccatcaatgg 840

```



```

gtgtcaggag ctgagcttgc cccgccagga acccctgagg gccgagcagc gcctcccagt 900
ggtcgggtact cacatgtagc tgcgggtgctt ggtggcagcg tcctgttggt ggctgggggg 960
tacagcggcc ggccccgtgg ggacttgatg gcgtacaagg tgccccctt tgtgttccag 1020
gcacctgccc ctgactacca cttggactac tgctccatgt acacagacca cagcgtctgc 1080
tcccggggacc cggaatgcag ttggtgccaa ggagcctgcc aagctgcacc cctcctggg 1140
accccccttg gggttgtcc agccgccagc tgctggggcc tgggcccgcct cctgggtgac 1200
tgccaggcct gcctggcctt cagcagcccc acagccccctc cacggggacc tggcacccctg 1260
ggctgggtgcg tgcacaatga gagctgcctc cctaggcctg agcaggcccc ctgccagggg 1320
gagcagatct caggcactgt gggctggtgg gggcctgcgc ctgtcttcgt cacgtccctg 1380
gaggcctgcg tcacccagag cttcctgcct ggctgcact tgctcacctt tcagcagccg 1440
cccaatacct cccagcctga caaggctca attgtccgca gcacgaccat caccctaaca 1500
cccagcgcag agacagatgt gtccctgggc taccgtggct tcacttacc aatgctgcct 1560
ggagggccag gtggaccagg ggctgaggac gtggccgtgt ggacgcgggc ccagcgccta 1620
cacgtcctgg cccggatggc ccgtggccct gacacggaga acatggtgag gccgcctggg 1680
acattcaggg ggttgtttat ggtaaataagg gagttagggg tgtggtatag agcatctgtg 1740
gtgatttgaa ggttctgggg aagggctggg gagaaattag aggcaggaga tggatttttt 1800
gtgtgtgtgt aagacagagt ctactctgt tgtccaggct ggagtgcagt ggcgcgatct 1860
cggctcactg ctaccttcac ctcccagggt caagcgattc ttctgcctca gcctcccacg 1920
tagctgggat tacagatgtg cattacacgt tggagcacat ggagctcagt gtgatcc 1977

```

&lt;210&gt; 85

&lt;211&gt; 1376

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3741930CB1

&lt;400&gt; 85

```

cccaactcct gccgcgcgcg gcctggcgtg caggaatgga gctcgggtccc tttaaatccc 60
gccgaggaac gcctgagggc cgggcccgcg cattcccgga tccgcgggct gggaatcgcg 120
ggtccccgcg cgtcctgcag cacttcccgc gccgcgcgcc cagcccagtg cgcgggtctt 180
gtgcccgggt gcccgggcgg tcgggagaaat gagccgcac taccacgacg gcgcgctccg 240
gaacaagggc gtgcagagcg tccggctgcc aggggcctgg gaccccgccg cccaccaggg 300
gggaaatggt gtcctgctgg agggagaaact gatcgatgta tctcggcaca gcactcttga 360
cactcatggc aggaaggagc gctactacgt gctgtatatc cggcccagtc acatccatcg 420
ccgtaaaattc gacgccaagg gaaatgaaat cgagcccaac ttcagcgcca ccaggaagg 480
gaacacgggc ttctcatgt cgtcctacaa ggtggaagcc aagggggaca ctgacaggct 540
cacgcccagc gcgtgaagg ggctggtcaa caagccagag ctgctcgcgc tgacagagag 600
cctcaccccc gaccacacag tggcgttctg gatgcccag tcagagatgg aggtgatgga 660
actcgagctg ggggcccggg tacggctgaa gactcggggc gatggtccct tcctggattc 720
attggccaaa cttgaggctg gaacagtgc caagtgtaat ttcactggtg atggaaagac 780
aggggcatcc tggacagaca acatcatggc ccaaaagtgt tcgaaggggg ctgcagcgga 840
gatccgagag caggggggatg gggcagagga cgaggagtgg gatgactgag ggacgtggct 900
gcctctgggg gcaccagcac cttctgtagc attgtggaac atttcttcca tgaacttcta 960
ctctggagct ggagagaaca gtctgccaag tcttacctgg ttcttcccca tagccacctt 1020
ggaatactct ttgaacaagc tctgctaaaa ccaaaccat gaccactgtg ttggtggtgg 1080
ggcctcatgc tgtggtcag ggagggcccc tccctgcatg tgagtctagc cttcctagac 1140
attacatgat gaccttttaa aaaaaaatta gagatggagg tctcactgtg ttgcccaggc 1200
tgatctcgaa ctctggact caagcgatcc acccacctcc acctcccaag gtgttgggat 1260
tgcaggcgtg agccactgca cccagcctca tgttgacctg gcctggatag cagctttgtt 1320
tgtagaggca ggtggtggga gagaccctag gacagcctga gagtaactga acctgg 1376

```

&lt;210&gt; 86

&lt;211&gt; 2635

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 5402506CB1

&lt;400&gt; 86

```

ccactgccct gtcgccccgc ctgtgcagcc caggacctag gcgcgggtagc tggggctggc 60
ttttgagggg cgcgggcagc cttctgactg ggtcggaggc ctgcggggccc gaagcctctg 120
tccctcctgt tcttgctcgg cgtgcttag cccctccgcg tagtcatcat ggatctgatt 180
ttaaacggaa tggattatct gcagggtggga gtaacatctc agaagactat gaagctaatt 240
cctgcctcaa gacacagagc tacacaaaag gtggttattg gagatcatga tggggtagtt 300
atgtgctttg gcatgaagaa aggagaagca gcagcagtg tcaagacttt acccgggccc 360
aagattgcaa ggctggaact gggagggggt atcaacacac ctcaggagaa aatttttatt 420
gctgcagcat ctgagattag aggccttcaca aaaagaggaa aacagttcct ctcctttgaa 480
acaaacctca ctgaaagcat taaagctatg cacatatctg gctcagacct ctttctcagt 540
gcaagttaca totataacca ttattgtgac tgcaaaagacc aacattatta cttttctggg 600
gataaaatca atgatgtgat ctgccttcca gtggaaagat tatctcgtat cacacctgta 660
ttggcctgcc aggacagagt gctcagagtt ttacagggat ctgatgtgat gtatgcagtt 720
gaagttcctg gacccccctac tgtcttagca ctacacaatg gaaatggcgg tgactctgga 780
gaagaccttt tgtttgggac atcagacgga aaacttgccg ttatacagat tactacatcc 840
aaaccagtac gcaagtggga aattcaaaat gagaaaaaga gaggaggtat tttgtgtatt 900
gacagctttg acattgtggg tgatgggggt aaagatttac ttgttgggag agatgacgga 960
atggttgaag tgtatagttt tgataatgca aatgaacctg ttctacgatt tgatcagatg 1020
ttgtctgaaa gcgtcacatc tatccagggt ggttggttag gaaaagacag ctatgatgaa 1080
atcgtggtgt ccacatatcc aggcctgggt acaggtctga caacagagcc cattcataag 1140
gaaagtggac caggagaaga actaaaaatt tgcaggaga tgcagaataa aatttcttcc 1200
ttacggaatg agttggaaca tttgcagtat aaggatttgc agggaaagaga gaattatcaa 1260
cagtcttctc aatcaagcaa agcaaaatca gcagtacct cctttggtat aaatgataaa 1320
tttactacta ataaagatga tgccagttac agccttatct tagagggtaca gactgcaata 1380
gataatgtct taatacagag tgatgttcca atagatttac ttgatgtgga taaaaattct 1440
gctgttgtaa gcttttagcag ctgtgattct gagtcaaacg acaacttctc tcttgccact 1500
tatcggtgcc aggcagatac tacaaggctg gaactcaaga ttcgctcaat tgaaggccag 1560
tatggcacac tacaagcata tgtgactcca agaattcaac ccaaaacctg tcagggtccg 1620
cagtaccaca tcaaacctct ttcactccat caaagaactc actttattga tcatgacaga 1680
cccatgaata cactgacct aacaggccag ttcagttttg ctgaagttca ctcctgggtg 1740
gttttttgtc tgccctgaagt tccagaaaaa cctccagcag gagaatgtgt gacattttac 1800
tttcagaaca cctttctaga tacacaactt gaaagtacct acagaaaagg agagggagtt 1860
tttaaatctg acaactttc tactatctcc atcctaaaag atgtgctttc taaagaagct 1920
acaaaaagga aaattaacct caacatatca tacgagataa atgaagtatc agtcaaacac 1980
actttaaagc taatccaccc aaagctggag taccagttgc ttttggctaa gaaagtgcag 2040
ttaattgatg ctttaaaaga attacagatt catgagggaa atacgaactt tctgatacca 2100
gaatatcact gtattctaga agaggcagat cacctacagg aagaatacaa aaagcaacct 2160
gcacatcttg aaagactcta tggcatgac actgatcttt tcatagataa atttaagttt 2220
aaaggcacca atgtaaaaac taaagtacct cttctatttg aaattctgga cagttatgac 2280
caaaatgcat tgatttcatt cttecatgca gcatgacata caaataagggt aaagtcccaa 2340
agaagtggtc tgtttaaatg aaaaatgtta aaagagaatt agaagttaac tgtgtatgta 2400
ctttaatttc aaatgctttt ttatatctaa aaatatttga gatataatgg ctttttaaaag 2460
taatgcttct atttcttttg acaaattgaa ctttctaaaa ctaaaagagt ctttttattt 2520
ttaaacaca agtagaatga tttaaatagg attttaatga atttttggca agtggctggt 2580
taaattttta aattgagaag taatatgtgg aaaaattaca tatatccaaa tgtttt 2635

```

&lt;210&gt; 87

&lt;211&gt; 3414

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 71081333CB1

&lt;400&gt; 87

```

gttcccagact ccggcaagag ggtcagtggt gaggcgggag cccggggcggc cggcttgaac 60
tccgtgcacg cgggcctgga gctgcgcccc tgcgtctttt cccgacttct gatccgcaga 120
ggggcactcg gagccccagc ctgggtcgca gtgccttcgg tctgcgggccc tcagtttctc 180
ggtggcgctt ggactctcgc cctcagaggg aggcctgctc caggtctgtt ggctacggat 240
ccaggagccc tcttcgctgg cgattcgcgt agcctcagat tgccctggat gatggcctgc 300
gactgacagt ctctcagctc ttctcagctc tcctgagget gcagctgtca gctccctgat 360
cctcagcact tgccctggacc agtgccgggtc tgccacgccc tgttcagcca tccctgtcct 420

```

```

gttgtgtcgt gcaagccaca tctgtccaca cccccaccac ccaccagtgc tggctctcct 480
ctcttctccc ctgtgtctccc agcatgcctc gtggcccggtg ccgtcagcag ggccctcgca 540
ttcccatctg ggcagccgcc aattatgcca acgcacatcc atggcagcag atggacaagg 600
cgtctccttg ggtggcatat accccccttg tggatccctg gattgagcgg cctctctgtg 660
gggacaccgt gtgtgtgcga acgaccatgg agcagaagag cacagctagt ggcaactgtg 720
gcggtaaacc tgcagaaagg ggtcccctag ctgggcacat gccagctcc cgaccccaca 780
gggtggactt ctgctgggta ccaggctcag acccaggcac ctttgatggc tccccgtggc 840
tactggaccg cttcttggcc cagctgggag attacatgtc cttccacttt gagcactatc 900
aggacaacat cagccgtgtc tgcgagatcc tcaggcgctt aacaggccga gcccaggcct 960
gggcagcccc ctaccttgat ggggacctgc ccctgcctga cgattacgag ctcttctgcc 1020
aggatctcaa ggaagtgtt caagaccga acagttttgc tgagtacat gctgtggtta 1080
cctgtcccct gccctggcc tccagccagc tgccagtggc ccctcagctg cctgtggtga 1140
ggcaatactt agctagggtt ttagaggggc tggcactcga catgggtact gccccaggt 1200
ctttaccagc cgccatggcc acccctgctg tgtctgggtc caactctgta tctagaagtg 1260
ctctgttcga gcagcagctg accaaggaga gcacccctgg gcccaaggag ccccagctcc 1320
tgcccagttc tacatgtagc tccaagcctg gtcctgtgga accagcctct tcccagccag 1380
aggaggcagc cccacacct gtccctagac tgtcggagtc agctaactct cctgccaga 1440
gaccagaccc agctcatcca ggaggccaac aacccagaa aacagaggag gaggtttttg 1500
agacagaggg agaccaggag gtgtccttag gtacccaca ggagggtgtg gaggccccgg 1560
agacccagg agagccacca ctttctcctg ggttctgaac acagcccagg tgggaacaat 1620
gctgccccct atgatgaagt ggcctgtgtg gcttgagcgc ccatagtcc ccagtcagag 1680
cagagtgggtg tcccagatg acttcagacc ccatagctgg gcaagatgcg cttgttttgg 1740
actctgcgct gagcagaacc agtccccca actccagcag agcttgacct ccgccctgtg 1800
ccctttccct gctgctggct ctctgctgca tccctgcccg tcttctggga gtgccgtctc 1860
accaggcct gctcccacga gggggtcgtt ttgtagatca actctcagca gatagttgca 1920
tcacttttgt caccctccacc cccataaaac accccccttg gtgtcttcca cactggctgg 1980
gactgaactg ggtctgccac gtctgccctg ttgggacctc gttcctgtat ggggtttctg 2040
tgtctgtccc cctagcccca ccctctctgc tccctgtgac ttctgtccca gctggaatgt 2100
gaggaaggct gcctggagga agtgggcact ttgattctgg gggatgagtg ggatagcatc 2160
ttggcaaaaga atgatgtttc ctgtgggatt accagcagtt ctgggcagct ggaaggtttg 2220
gatgcaggaa ggtgggatgc aagccccttg ccaggctcag ctgcaggcat tcgccctttg 2280
tgtgttaaca ctgtagctct aatcgtgcct cgtgttctga agcaggctcc tcttccctgc 2340
tgctgtgctg ccatccccctg ctactccgc actcgcagtc attgcggatt cagcacttaa 2400
gccctagtgt gtaaatgact tggatctccc tcctggcaga ctgtgacct cagcacaggg 2460
gctctgtctt agtcagctct gtggcctggg acacaattgg tgctcattaa acacgtgttg 2520
gattactggc tgagggtggg cagtggctgc ctaccagcaa gtctttgttt cccattctgt 2580
caggagcaga ggcctggcca tgagcattcc aggtcagcat tgcagctgtg gtttctggca 2640
ggggccttga ggggtccaagg gtctgatcaa agaagctggg tggggctgag tagggggtgg 2700
gtggctctgc aggtagagaa ctgacttccg agagctgtag gtgaagtgag gaccaggcag 2760
cagtcagag ctgtgaggcc ccaggccag aggaatggaa tgaagaaaga cctgttccac 2820
acaaggaggg gttttctagt ggaagctgag cttggaagct cctggagcca agtttaccac 2880
aggataggca ggtacagccg cccagtgtgt gcagtgtgtg cccactctgg tgcttgaca 2940
agaagttcaa gtccctttca ggtttacaat tgtgaaaaa acaaaaacaac acctcctgg 3000
aaagtaggaa gcacaaagcag aaatgtgaat tggaaaata gccaactatt ttaggaaaaa 3060
ctaattagca agcatctgat taaacttatt cagaaaactg ccgggagcag tgtctcacac 3120
ctgtaatccc agcactctgg gagccaagg tgggtggatc acctgaggtc aggagttcaa 3180
gaccagcctg accagcatgg agaaatccat ctctcctaaa aatacaaaat tagccaagt 3240
tggtgggtga tgctgtaat ctgagctact tgggaggctg aggcagaaga attgcttga 3300
cccaggaggc agaggggtgt gtgagccgag atcacgccat tgcactccag cctggacaac 3360
aagagtga aa ctccatccca aaaaaaaaaa aaaaaaaaaa aaaatggtgc ggtc 3414

```

&lt;210&gt; 88

&lt;211&gt; 936

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7503139CB1

&lt;400&gt; 88

```

gggcacccca agagcaggcc tggagacccg ggagtgcctt ggggagaggg tccccgccct 60
gccgggagcc cggccctccc ggtgccgggg tggacactgt tagtgtagtc attgtcccta 120

```

tggagaaact	gagggcatgga	gaggctgaag	tcacacaagc	ccgccacgat	gactatcatg	180
gtggaggaca	tcatgaagct	gctgtgctcc	ctttctgggg	agagggaagat	gaaggcggct	240
gtcaagcact	ctgggaaggg	tgccttggtc	acaggggcca	tggccttcgt	cgggggtttg	300
gtgggcggcc	caccgggact	cgccgttggtc	ggggctgtcg	gggggctggt	aggtgcctgg	360
atgacaagtg	gacagtttaa	gccggttcct	cagatcctaa	tggagctgcc	ccctgccgag	420
caacagaggc	tctttaacga	agccgcagcc	atcatcaggc	acctggagtg	gacggacgcc	480
gtgcagctga	ctgcgctgg	catgggcagc	gaggccctgc	agcagcagct	gctggccatg	540
ctgggtgaact	acgtcaccaa	ggagctgcgg	gccgagatcc	agtatgatga	ctaggccgca	600
cctccggggg	ggtggggggc	cccctttaaa	tgactctgtg	attctgaaga	ggtggcttgg	660
gagttggggag	aagccccgcc	ggatgcccc	tggggaatct	ccacatcatc	cagtgtatta	720
ctagtaaatgt	ccccgctgga	gaggcccccc	gctgtgccag	tgtccatggt	cccagaaaatt	780
actgatgaag	ccagccatgt	gttggtggca	tgtgccctgg	cctgccatga	cagccctctg	840
actggcccc	ccagtgaaga	gtaaaggcct	gcctgccgcc	ggcttcggag	gcgtctgctg	900
agtcctctca	cccgcattgg	tctggggaag	tgatca			936

&lt;210&gt; 89

&lt;211&gt; 3123

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7505836CB1

&lt;400&gt; 89

aagaagctaa	gatagccaac	tgattataag	gagaccatct	gaccacaatt	ttctactact	60
gcagaagcat	tgctgtgaag	ttccctttcc	cagctgcctc	cactaatctg	caaaaagcac	120
tttctaaagc	actggaaaagc	cgagatgagg	tgaaaaccaa	gtgggggtgt	tctgacttca	180
tcaaggcctt	tattaaattc	cacggtcatg	tgtacctgag	taagagcttg	gaaaagtgtg	240
gccctcttcg	agagaaaattg	gaagaacagt	ttaagaggct	gctattccaa	aaagctttca	300
actctcagca	gtaggttcat	gtcactgtca	ttaacctgtt	tcaacttcat	caccttcgtg	360
acttttagcaa	tgaaaccgag	cagcacactt	atagccaaga	tgagcagcta	tgttggacac	420
agttgtctggc	cctctttatg	tcttttcttg	gcacccctgt	caagtgtcct	ctacagaatg	480
agtctcagga	ggagtcctac	aatgcctatc	ctcttcacgc	agtcaaggct	tccatggact	540
ggctaagact	cagacccagg	gtctttcagg	aggcagtggt	ggatgaaaga	cagtacattt	600
ggccctgggt	gattttctct	ctgaatatgt	tccatcccca	tgaagaggac	ctctcaagta	660
ttagtgcgac	accacttcca	gaggagtttg	aattacaagg	atttttggca	ttgagacctt	720
ctttcaggaa	cttgggattt	tccaaaggtt	accagggtat	tacaggggac	aaagaaggcc	780
agcaacgacg	aatacgacag	caacgcttga	tctctatagg	caaattggatt	gctgataatc	840
agccaaggct	gattcagtgt	gaaaatgagg	tagggaaatt	gttggtttatc	acagaaatcc	900
cagaattaat	actggaagac	cccagtgaag	ccaaagagaa	cctcattctg	caagaaacat	960
ctgtgataga	gtcgtggct	gcagatggga	gcccagggtc	aaaatcagtg	ctatctacaa	1020
gccgaaattt	aagcaacaac	tgtgacacag	gagagaagcc	agtgggttacc	ttcaaaggaaa	1080
acatttaagac	acgagaagtg	aacagagacc	aaggaaagag	ttttcctccc	aaagaggtaa	1140
aatcccagac	agaactaaga	aagactccag	tgtctgaagc	cagaaaaaca	cctgtaactc	1200
aaaccccaac	tcaagcaagt	aactcccagt	tcatcccat	tcatcaccct	ggagccttcc	1260
ctcctcttcc	cagcaggcca	gggtttccgc	ccccaacata	tgttatcccc	ccgcctgtgg	1320
cattttctat	gggtcaggt	tacaccttcc	cagctgggtg	ttctgtccca	ggaaccttcc	1380
ttcagcctac	agctcactct	ccagcaggaa	accaggtgca	agctgggaaa	cagtcaccaca	1440
ttccttacag	ccagcaacgg	ccctctggac	cagggccaat	gaaccaggga	cctcaacaat	1500
cacagccacc	ttcccagcaa	ccccttacat	ctttaccagc	tcagccaaca	gcacagtcta	1560
caagccagct	gcaggttcaa	gctctaactc	agcaacaaca	atcccttaca	aaagctgtgc	1620
cggcttttggg	gaaaagcccg	cctcaccact	ctggattcca	gcagtatcaa	caggcagatg	1680
cctccaaaca	gctgtggaat	ccccctcagg	ttcaaggccc	attagggaaa	attatgcctg	1740
tgaaacagcc	ctactacctt	cagacccaag	accccataaa	actgttttag	ccgtctattg	1800
aacctcctgt	aatgcagcag	cagcctctaa	aaaaaaaaat	gaagcctttt	cccatggagc	1860
catataacca	taatccctca	gaagtcaagg	tcccagaatt	ctactgggat	tcttcctaca	1920
gcattggctga	taacagatct	gtaatggcac	agcaagcaaa	catagaccgc	aggggcaaac	1980
ggtcaccagg	aatcttccgt	ccagagcagg	atcctgtacc	cagaatgccg	tttgagaaat	2040
ccttatttga	gaagccctca	gagctcatgt	cacattcatc	ctctttcctg	tccctcaccg	2100
gattctctct	caatcaggaa	agatacccaa	ataatagtat	gttcaatgag	gtatatggga	2160
aaaacctgac	atccagctcc	aaagcagaac	tcagtccctc	aatggccccc	caggaaacat	2220
ctctgtattc	cctttttgaa	gggactccgt	ggtctccatc	acttctctgc	agttcagatc	2280

```

attcaacacc agccagccag tctcctcatt cctctaacc aagcagccta cccagctctc 2340
ctccaacaca caaccataat tctgttccat tctccaattt tggacccatt gggactccag 2400
ataacagggg tagaaggact gcagatcggt ggaaaactga taagccagcc atgggtgggt 2460
ttggcattga ttatctctca gcaacgtcat cctctgagag cagttggcat caggccagca 2520
ctccgagtgg cacctggaca ggccatggcc cttccatgga ggattcctct gctgtcctca 2580
tggaaagcct aaagtctatc tggccagtt ccatgatgca tcctggacct tctgctctgg 2640
agcagctgtt aatgcagcag aagcagaaac agcaacgggg acaaggcacc atgaaccctc 2700
cacactgagg ccaaagtggc aacctgggaa tgaaggctcc ataaaccatg gcatgttggg 2760
tttgaggagc tggcccacac agtcccctgc aggtggcagc cctcttttct gtttctcgct 2820
gtcaagaggg tgtaagtatt ccaccagccc gctgagtgtg cacgaaatgt tcgcagtgtg 2880
acaaaaagaa aaatccatca ggaactctcc gtcccccccg ggccttcccg agggagagag 2940
agaggaactg ctgtttatct cactcagtta cttggtatca ccgcctctca ctttctccat 3000
cgtgcatgtc cccagccaca tgggaagtga aagctgagaa gggaaaggcag atgggagaa 3060
ccaatgggaa cttctcagtc cttttttcct ctttggggaa taaaatagga atccattaaa 3120
aaa 3123

```

<210> 90

<211> 616

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7505858CB1

<400> 90

```

agtcgcgacc ctggtccgga cctgacctga attgcgaccc caacctggac tgctcccctg 60
accgcaaccc ctacccccgc ccaccagtat ggcccggcac gtgttcctaa cggggccccc 120
aggagtggga aaaacaacat tgatccataa agccagttag gttttaaaat cctctggtgt 180
gcctgttgat ggattttata ccgaagaagt cagacagggg gggagaagaa taggattcga 240
tgtcgtcacg ttgtccggca cccggggggc tttatcgaga gttgggttag agcctccacc 300
tggaaaacgt gaatgccgag ttgggcagta tgtggtcgac ctgacttctt ttgagcagtt 360
ggcactaccc gtcttgagga atgtcaccaa ggaaaacaga aaccaccttc tgccagatat 420
cgtgacgtgc gtgcagagca gcaggaagtg aagacacgtg cattcctgcc ttccgtgaag 480
gagtgccccag ttcaagagga gcctgatgga gccctgcctg tcgaggctgt atgcctatgg 540
ggttatggaa ccttgtgggc ttttctagag aaaactcaac agctgtttcc cataaaatgt 600
ttaaagatc aaatta 616

```

<210> 91

<211> 1834

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7505872CB1

<400> 91

```

gcgcccagca agccccggaa ggcaggcgcg cgggttagaa cgcgccagag gtcggcgcg 60
gcacacccgc accgccccga cccaggttag tgaggccagt gattccgagt gtgtgaggag 120
cggcagtggc ggcgaggag gggcgggcgt ggggtggggc gggggcgga tgcgtcccc 180
ggcccctcta gcccgtggt ggtacaacgc gaaggtgtgg gaaggccgcg ataaaccgga 240
actgcagccc gccggacacc tccggcttca cttccgtaag aggagaggag tgtacggcaa 300
ggggcgggaa ctggaacttg gccgcctcg ttgtgagctg agctcagcgg cacgcttttg 360
tggcgtaact gcactgttac cccgcctac gtgtctctga cgctgacacc ttctcactgt 420
gaaacgtcgc gacctgtgac gtctgggggg cgcctcaaat cttccactcc agcatcggat 480
cccggaaagg cagcgtcgga gactggaccc aaaactcttc ctgttctgcc tgcagagttg 540
agccccgtcc gggtcctgga cccgcgtagt actgaccctg gatccctgtt cactgcgttc 600
tcgctccccg cgctccctgc tggaccccg gatgccgggc atctccgcc gaggcctctc 660
tcattgaggg aggaagcagc tagctgttaa cctcaccgtg gtcctggcac tctaccgttc 720
catcttggat gcctacatca tcgaattttt cacagacaac ctatgggaca cactcccttg 780
ctcatggcag gaagcattgg atggactgaa accaccacag ctggccacaa tgctgctggg 840
gatgcctggg gaaggggagg tcgtcaggta caggtcagtg tggccactca ccctgctggc 900

```

```

cctgaagtcc acggcgctgtg ccttggcctt tacccggtatg cctggctttc agacccccctc 960
agaattcctg gagaaccccc gccagagctc ccgactaaca gctccattcc ggaaacatgt 1020
caggcccaag aagcagcatg agatccggag gctgggagag ttggtgaaga agctgagtga 1080
tttcacaggc tgcacccagg ttgtagacgt gggctcaggc cagggccatc tctcccgtt 1140
catggctctt ggcctggggg tgatggtgaa gagcatcgaa ggggatcaga gactggtgga 1200
gagagccag cgctggacc aggagcttct gcaggctctg gagaaagagg agaagaggaa 1260
cccgcaggtg gtccaaacca gccctcgtca ctccccacac cacgtgggta gatatgtgca 1320
gcgggggcta cagcgagtgg ggctagatcc ccagctgcca ctgaatctgg ctgcccttca 1380
ggcccacctg gcccaggaga accgtgtggt ggccttcttc agcctggctc tactgcttgc 1440
cccactgggt gagacgttta ttctactgga ccggctgctg taccttcagg aacagggttt 1500
ccatgctgag ctctgcccc tcttcagtcc tgaactctct cccagaaacc tggttctggt 1560
ggccaccaag atgcccctgg gtcaggctct ttctgttctg gagactgaag acagctgatg 1620
cagcctgagg agacatctca gaccccatca tctgaaagtg cccagagagc acagtggcag 1680
agtacatctc atccagagaa acagcatcct gcactctcca gagtcttggt tccttcagtt 1740
tcatccctt tctctcctt catggattat gtaatacatt gtaaagtttt aattaattaa 1800
aaattggata tctgaaaaaa aaaaaaaaaa aaaa 1834

```

```

<210> 92
<211> 1340
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7506456CB1

```

```

<400> 92
cccaactcct gccgcgcgcg gcctggcgtg caggaatgga gctcggtccc tttaaatccc 60
gccaggaac gcctgagggc cgggccgcgc cattcccga tccgccggct gggaaatcgcg 120
ggcccccgcg cgtcctgcag cacttcccgc gccgcgcccc cagcccagtg cgccggctct 180
gtgccggggg gcccgggcgg tcggggagaat gagccgcac taccacgacg gcgcgctccg 240
gaacaaggcg gtgcagagcg tccggctgcc aggggccttg gaccccgccg cccaccaggg 300
gggaaatggg gtcctgctgg agggagaact gatcgatgta tctcggcaca gcatcttgga 360
cactcatggc aggaaggagc gctactacgt gctgtatatc cggcccagtc acatccatcg 420
ccgtaaatc gacgccaagg gaaatgaaat cgagcccaac ttcagcgcca ccaggaagg 480
gaacacgggc ttcctcatgt cgtcctacaa ggtggaagcc aaggggctgg tcaacaagcc 540
agagctgctc gcgctgacag agagcctcac ccccgaccac acagtggcgt tctggatgcc 600
cgagtcagag atggaggtga tggaaactga gctggggggc ggggtacggc tgaagactcg 660
gggcgatggg ccttctctgg attcattggc caaacttgag gctggaacag tgaccaagt 720
taatttact ggtgatggaa agacaggggc atcctggaca gacaacatca tggcccaaaa 780
gtgttcgaag ggggctgcag cggagatccg agagcagggg gatggggcag aggacgagga 840
gtgggatgac tgagggacgt ggctgcctct gggggcacca gcaccttctg tagcattgtg 900
gaacatttct tccatgaact tctactctgg agctggagag aacagtctgc caagtcttac 960
ctggttcttc cccatagcca ccttgggaata ctcttgaac aagctctgct aaaaccaaac 1020
ccatgaccac tgtgttggtg gtggggcctc atgctgtggc tcagggaggg cccctccctg 1080
catgtgagtc tagccttct agacattaca tgatgacct ttaaaaaaaaa attagagatg 1140
gaggtctcac tgtgttgccc aggtgatct cgaactcctg gactcaggcg atccaccccc 1200
ctccacctcc caggggtgtt ggattgcagg cgtgagccac cgcaccacgc ctcatgttga 1260
cctggcctgg atagcagct tgtttgtaga ggcaggtggg gggagagacc ctaggacagc 1320
ctgagagtaa ctgaacctgg 1340

```

```

<210> 93
<211> 7706
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7506697CB1

```

```

<400> 93
ggtgtctgag atagcctcgt tgaatgtgta gggttcgagc ctggagttag gccagattgt 60
gtcgttttac ttgccttggg cgtggagaac gatcttgtga gaatatcttc aaaggcagaa 120

```

```

aatatttccc tttatgaatt ctcttttcccc ctgcgtgtaa gtcgggaatg tgaagaggag 180
tgtaggaaag agccctgggt caagtaggtc ttcagaacta ctagcagata atttggggggg 240
acttcatatt cagaagtcta tataaaggat tcctcatgtc cataacatgt tggatgagggc 300
tctgcagctc accccactc tcagagtggg cagctcccat taattggacc ccgtgatttc 360
caatctctgc tgtgttggac gtcattgagca ttgcaatccc tctggggagtc accacatcag 420
atacatccta ctcagatatg gctgctggat cagaccctga atctgtggag gctagtccag 480
cagttaatga gaagagcgtg tattccactc ataattatgg gaccactcag aggcattgggt 540
gtcggaggact gccttatgct acgatcatcc ctgcgttctga cctgaatggc ctgccgtcgc 600
ctgtagagga acgctgtgga gacagcccga actctgaagg agaaactgta cctacctggg 660
gtccttgtgg tctttctcag gatggcttcc ttctcaactg tgacaagtgc aggggaatga 720
gcaggggggaa ggttattaga cttcatcggc ggaagcagga caacatatca ggtgggggata 780
gcagtgcac acctacaagc atcaccttaa ctggttagaag aaccaaacc cagaagcggg 900
cacagcacac aaaaagggt cgtgcagcac caaagacgaa gaaaatcaag aattctccct 960
ctgaagcaca gaatttagat gagaatacaa ctgagggctg ggaaaatcgg ataagactat 1020
ggactgacca gtatgaagaa gctttcacta atcagtagag tgcagatgta cagaacggcg 1080
ttgaacaaca cctacattct agcaaggaa ttgtgggcaa acctactatt ttagacacta 1140
ttaataagac tgaattggcc tgtaataaca cagtatttgg ttcccaaag cagttacagc 1200
tggaagaggt cactcgtgtt caaagcacc ggaagatcct gagggctgca agagatttgg 1260
ctttggacac tcttataata gagtatcgtg ggaagtcac gttacgacag caatttgagg 1320
tcaatgggca tttcttcaaa aaaccatacc ctttgtgct cttctactca aaattcaatg 1380
gtgtagagat gtgtgtggat gcccgactt tcggtaatga tgctcgggtc atcagaagat 1440
catgtacacc aaatgcagag gtgcgacaca tgattgcaga tgggatgatt cacctgtgca 1500
tctatgctgt gtctgccatc accaaggatg ctgaggtcac catagcattt gattatgagt 1560
atagtaactg taattataaa gtggactgtg cctgtcacaa gggaaaccgg aattgtccta 1620
tacaaaaaag gaatcctaag gctacagaac tgccactcct accactcct ccaagcctac 1680
ccaccattgg agcagagact agacgtagaa aagcacgacg gaaagagcta gagatggagc 1740
agcagaatga ggcttcagag gagaataatg accagcaatc acaagaagt ccagaaaaag 1800
taactgtatc cagtgtatc gaggaagtag acaatccaga agaaaaacca gaagaagaga 1860
aagaagaggt tatagatgac caggagaacc tagctcatag caggaggacc aggggaagata 1920
gaaaggtaga agccatcatg catgcttttg aaaacttaga gaaaagaaag aagcggcggg 1980
atcagccctt ggaacagagc aactctgatg tagagattac tacaaccacc tcagagactc 2040
ctgttgggga agagacaaaa actgaagccc ctgaatctga agttagcaac tctgtttcaa 2100
atgttaccat cccaagcacc ccacagagtg ttggtgtgaa taccggaggg tcttcccaag 2160
caggggatat tgctgcagaa aaactagtcc ccaagccacc tccagcaaag ctttctaggc 2220
cccggccgaa gagtcgaatt tctcgggtaca ggaccagttc agcccaaaga ctaaagcgtc 2280
agaagcaggg caatgcacag caggcagaat tgtcacaagc tgccttgga gagggaggaa 2340
gtaacagttt agtaactcct actgaagctg gaagtctaga cagttcagga gaaaacaggc 2400
cattaacagg gtctgaccca actgtggtgt aacttactgg atcccatgtc aaccgtgtc 2460
catctaaata ccccaaaacc aaaaagtatc tagttacaga atggttgaat gacaaagcag 2520
agaagcaaga gtgccctgtt gagtgccctt tacgtatcac aacggatcca actgtactgg 2580
caacgaccct aaacatgtta ccagggtctta tccattcccc gttaatttgc accaccccca 2640
aacactacat tcgcttttggc tcacccttta tccctgagag acgtcgaagg ccccttctgc 2700
ctgttgggca agtagacaaa tghtaagaag cctggataaa acaagcctta gaagaaggga 2760
tgactcaaac atcatctgtt cccaagaga ctagaactca gcacctatac caaagcaatg 2820
agaatagtag ctcttctagt atctgcaaag acaatgcaga cttgttgagc cattaaaga 2880
aatggaagtc tgcgtatctg atggagcaga atgtcaccaa gttacttcgg cctctgtctc 2940
cagtcacacc accccctccc aattcaggct caaagagtcc ccagctggcc acacctggct 3000
catctcacc aggagaagag gagtgtcgaa atggatacag cctcatgttt tcaccagtca 3060
catctcttac tactgtagt cgtgtcaaca ctccctaca gtttgagctt tgtcaccgaa 3120
aagacctgga tttggcaaaa gtaggatacc ttgactcaa cactaacagc tgtgctgata 3180
gaccttccct actcaactca ggtcattctg acctggctcc tcatccctcc ctgggacca 3240
cttctgagac tggtttccca agcagaagtg gagatggaca tcagaccctc gtgagaaact 3300
cagaccaggg atttcggaca gagtccaact tgatgtatgc ctactccctt ttgaatgcta 3360
tgccctcagc agatggactg tatcgaggat ctctctagt gggggatagg aagcctttac 3420
atgtggatgg gggatattgt tcccctgcag aaggattttc cagcagatat gaacatggct 3480
taatgaaaga cctctctcgt ggatccttgt cacctgggtg tgaaagggcc tgtgaaggag 3540
tcccatctgc ccccagaac ccaccacaga ggaaaaaagt atccctgctg gagtaccgaa 3600
aacggaaaca agaagctaag gaaaattctg ctggtggggg aggtgactct gcacagagca 3660
aaagcaagtc tgcaggagct gggcaaggca gcagtaactc cgtttccgac actggtgccc 3720
atggtgtgca gggatcctca gccggaactc catcttcccc tcacaaaaaa tcttccccat 3780
ctcattctc tatgtcccat ttggaggcgg taagccatc agattccaga ggcacttctt 3840
catctcactg cagacctcaa gagaatatca gcagtaggtg gatggttccc acatcagtag 3900

```

aacgactccg	agaaggaggg	agcatcccca	aggctcctccg	aagcagcgtg	agggtggccc	3960
aaaagggaga	gccctctccc	acatgggaga	gtaacatcac	agagaaagac	tcagaccctg	4020
cagatggaga	aggcccagag	acattaagct	cagcactctc	taaaggagca	acagtttaca	4080
gcccttccag	atacagctac	cagctcctgc	agtgtgatag	tcctcggaca	gaatcacaaa	4140
gcctccttca	gcagagtccc	tcccccttca	gaggacatcc	tacacagtct	ccaggataca	4200
gttatcgaa	tactgcactg	agacctggaa	acccccctc	tcacggttct	tcagaatcat	4260
ccctctcttc	cacgtccctat	tccagccccg	cccaccctgt	gtccacagac	tcgttggccc	4320
catttacggg	gacaccaggg	tatttttagca	gccagccaca	ttctggaaac	agcactggca	4380
gcaatcttcc	aaggaggagc	tgcccttcta	gtgctgctag	ccctaccctg	cagggaccct	4440
cagactcgcc	aacctcagat	tcagttttctc	agtcacagac	aggaactctg	agttccacct	4500
cctttcctca	gaactctagg	tcgtcattgc	catcagactt	acggactatc	agtctgcccc	4560
gtgctgggca	gtcagctgtc	taccaggcct	ccagggtatc	tgcggtttcc	aattcacagc	4620
actaccacac	ccgtgggagt	gggggtgtgc	accagtaccg	actccagcca	ctgcaagggt	4680
caggagtcaa	gactcagacg	ggacttttct	agggttctctg	gatttgggca	aacagaactg	4740
aatgagccca	tagctgcttc	cttccagctg	cctctgggaa	ctaggccgag	catattgctg	4800
aggaacgggg	ggtacaaggt	gccagaggat	tgggtctggt	ggacaagaaa	caagacttgt	4860
ggtcacaatt	ggcctctggc	cttggagaaa	gctgtaaatc	ttgtctgaag	cagagactat	4920
aaagaagtgt	ctccctgctg	tcaagggtac	attgttgaca	agcaaatggt	gtttcgggta	4980
gtaacggttc	taagtgcatt	gagttgtgtt	gaagcctccg	tctcccatcc	ttgcctgtag	5040
cccgtagtca	cttgtgcagt	gaggacatct	ttttaaat	aaaaaaaaaa	aaaaaaaaaa	5100
ttttcaaaaa	aaaaaaagtt	aaaagagcca	atctcaaaagc	cccaagccat	ctgagtactg	5160
ttagggtttt	atgcacttaa	gaaaaaaggt	aggatatgtaa	atgttcatcc	taagacaacc	5220
attccaaaag	caggatctctg	gccaatgtgt	gtccaccaag	aatactgttt	atctttgtct	5280
taagatcacc	aagaaatagg	caaggatagt	aaagcttgga	acctgcacca	actggagggt	5340
gcctggctct	ttgaagaaaa	gctcatggtc	agctcttgat	tattcgggag	cagattat	5400
gagtagattg	tctgagcctc	caactgttac	catcctactc	ccccttccca	agctatttca	5460
cagctcagta	acccatgaag	taagtagaca	agaaaaggag	gaatgagaca	tgatataggc	5520
caattgcatt	gctacttacc	agcttttggc	aataaat	attaggaagg	ataccgagtg	5580
gtttgggaat	gcttcgaatt	ttattttttc	tactcccaat	taatcaggag	ttgatgatcc	5640
catgagcagg	accgctcca	tgattgggga	gcatgcactt	gtgactgcag	ggtaagagtg	5700
ggaagatagg	tttgtggagt	ggcaccgaca	ggactgtgat	tgtgtgtggg	cctgccccac	5760
atttctctag	gggatgctta	tgtgagagtg	ggcccagtg	aagagttacc	aagccacca	5820
caccctaac	actgttctgg	atgagagatg	agagcagacc	ggcttctccc	catcagtgca	5880
ttgtgctgt	tgtacacccc	tggaggagcc	ctggagccag	cccagggtggg	gtacacaatc	5940
tttttaaatt	ccatatgggt	gccagcttat	ttctttcact	tgtttactgt	aatatctggc	6000
gtgtttttat	ttatctaatt	ttgtattcag	ttataaccat	ggtaggggta	gtgaatat	6060
gacaggtgta	atccctgggtg	ctgcagtgga	ccttcttttc	ttttggacaa	gataatactg	6120
tgagtttccc	tccttctctc	cctctaattt	gttttctctt	tttcccagc	ctctgctatc	6180
cccttctttt	ctaccctgtc	ctacaactat	catatgcaca	gtcttctctc	tttgtgtgtg	6240
actgttacaa	aatttcactt	ttcaaaatcg	aaatcagggtg	tttgctcaaa	tgaggggaga	6300
tttttttttt	tttttttttt	taaatgctga	gacctcagca	gagtactttt	ctttttgttg	6360
tttccccac	aaacccatca	gtctgggaga	gcattgggag	tggaaatcat	gttgccctggg	6420
atgctggttt	ctttgtatat	tatataaaac	gtatgtaaat	gtctctccat	ttgggctggg	6480
gtttgcattc	tccccttggc	tatttaacca	aggggagagg	ccagcgggca	ggcggccctc	6540
accctgcctg	gcacgtgcag	agaccccagc	cactctgtgt	gggcagggtg	ctgtcaagac	6600
cagacctctt	gggggggtag	gggcgggggg	gtggggggaa	ctcttggaag	ggaagaagta	6660
tcacttcttt	ctcaagtgga	gtgtttacac	cttgctgtaa	catttgaaact	ttcacaagag	6720
atgtaataat	tttgataata	aaattcttaa	ccataataat	cataaagttg	agaacttctt	6780
tgtccaattt	tgtggccctg	ccttgcctgg	ttttccta	gccattttta	tctgaggaga	6840
atgtgccctt	gggtggagttg	ttttgcccat	gataatgggg	ccaggaatac	tgtctgatct	6900
tccttcgctt	tttatgattg	cagtcaggac	agtgtgaagg	caaacctgca	gcagggcaag	6960
tgaacattaa	aaactgcctt	gaacttgagt	gggtgtcctg	accttggcct	cattcattca	7020
gcaactat	attgggtctt	ttcattcaaa	aacaatgctg	gacattgttt	ttgggtgctg	7080
gcttatgggt	attagaatgc	tgattggggg	tgggcaaagt	aaggtctctt	ctttcatgga	7140
gtttacattc	tagtggagga	gagaactaaa	ctagtata	attttaaata	gtcattgcca	7200
tgtaattgt	ttcatagcag	atatgtgact	aagacagttg	aaggagccac	tctagattgg	7260
ttggtaaga	aaagtctctg	aggagtgtac	atttaaactg	agacttcagt	gacaaggagg	7320
cagtcctggg	acgatgtgag	tgaaattcat	tcttgaacgt	tgcagactct	gtggtagaaa	7380
aagcttgggt	aagatgtaat	aaatggatgg	ggctgggaaa	gaggatagaa	gatcctgtaa	7440
gagggagata	ggggagagat	catgtagagc	aggagtcggc	aaactctttc	tataaagggc	7500
cagatagtaa	atatttagac	tttgcaagcc	atgtatgggtc	tctgtcataa	ctattcagcc	7560
ctttggctgg	gtgcgaatca	tgctgtaat	tccagcactt	tgggaggccc	aagagagcag	7620
atcacttgag	gtcagaagtt	cgagaccagc	ctggccaaca	tggcaaaacc	ctgtctacta	7680



aaaatacaaaa aattaaccag gtgtgg

7706

&lt;210&gt; 94

&lt;211&gt; 1663

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7623472CB1

&lt;400&gt; 94

```

gtttggggcgc gaattactct aagcgcgggct ctcagaaggg tgcaagaaga atcagtgatt 60
tttttttttt tcgaagcact gtgttttagat caagatgtct ggtaaagcaa atgcttccaa 120
gaaaaacgct caacagttaa aaagaaatcc aaagagaaaa aaggataatg aggaagttgt 180
gttgtcagag aataagggtta gaaacacagt gaaaaaaaaat aaaaatcatc tgaaagatct 240
gtcttctgaa ggacaaacaa agcacactaa cctaaaacac ggaaagacag cagccagcaa 300
gagaaaaacc tggcaacctc tgtcaaagag taccagagac catttgcaaa ctatgatgga 360
atcagtaata atgacaattt tgagtaacag tattaagaa aaagaagaaa tacaatacca 420
tctcaacttc ctgaagaaaa gattgctaca acagtgtgaa actctgaaag tccctcccaa 480
aaagatggaa gatttaacta atgtatcaag tctactgaat atggaaaagg cagcagacaa 540
agctaataaa gaagggtctgg cattactaca ggaagaaata gataaaatgg tagagaccac 600
agagttaatg actgggaata ttcagagcct aaagaacaaa attcagattc tggcaagtga 660
ggtggaagaa gaagaggaga gagtaaaaca gatgcatcaa ataaatagta gtggagtact 720
ctctcttcgc gaactttctc agaaaactct caaagcacc acacttcaga aagaaatatt 780
ggcgctaatt ccaaaccaga atgctcttct aaaggacttg gatattcttc ataattcatc 840
acagatgaag agcatgtcaa ccttcattga agaagcctat aagaaactgg atgcatctta 900
aagagtgttt tttttttaga ttgttccata ttaatttaat gttcgtgaat ttgtaaaact 960
gttaacctat gattatatgt acagaggcta aggtctctgc aggatttatt atctcctgat 1020
atgcacttta aaattagtct ttgtagtctc atcattagca tctaattgat ttctgaagac 1080
tgttttttagc agttgcaaaa tctaggaaac taacatttat attgctgtat ctataagtat 1140
ttgcctaaaa ccataaaaa ggaattgcct taaggatcta atagtttata atgcctgaat 1200
ggaactcatt tattgggtta cagctcatat caggagctgt gataagtaag tgtttacata 1260
cactgtctag tgcataaaaa ccttagtggt ttgctatcatt cattctgtag ccaatggcct 1320
acagtttctg tctgggtcatc tggaaacttg aaaatcctca aatgccttca ctcttttacc 1380
taactcccta gtcttttttg gaagtagtta agatatgcct gctatcttta actttggagg 1440
tagtataaaa atatttcttc tttggttggt tagtgaatta atgttgaata ttttctactg 1500
tcataaaaac atattcccaa aaagtattaa aaaggaacct tagcacattt attttattaa 1560
aacaagaaaa cccaaaatat ataaaccttt tcttattatg ataataagct tatatatgtg 1620
aataagttat aacctagtgt ataaaaaatt atttctaaca aaa 1663

```

&lt;210&gt; 95

&lt;211&gt; 2144

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7506416CB1

&lt;400&gt; 95

```

ccggaagtgc ccgcgcaccg gtttccgggt caggtgggga aaatggcggg gtctacagtg 60
ttctcgactt cgtcgctggg aattgggttg tttgttgctg ttggctcagt gtctccatcg 120
gggaaggagg atctgggggt tctactcca ccactttcac ctcttgctga ggtcaccacc 180
ttgacatcat atcttgataa gttgtccact tctctggatt gggagccttt gggggattag 240
ggatcatgct gtctctgtgt gctcaggatt tgacacagat gcttgctctg tcaaggcaca 300
gcctattgtc tctttgtctc agtgtgacat cattcagacg cttctacaga ggtgacagcc 360
caacagattc ccaaaaggac atgattgaaa tccctttgcc tccatggcag gagagaactg 420
atgaatccat agaaaccaa agagcccgcc tgccttatga gagcagaaag aggggaatgt 480
tggaactcgc cattttctt agtctttttg ctaaagaaca tctgcagcac atgacagaaa 540
agcagctgaa cctctatgac cgctgatta acgagcctag taatgactgg gatatttact 600
actgggcccac agggagaagg ttttatacaa gaaaatggca cattctgaag tggagctcaa 660
cagacagcaa ctcttcccag ccctgtggag gaggccggag gttggggccg gagccctgga 720

```

```

agcagggcct tgctcgggca gcttctgacc ctccccctct ggccaggccc cctggggccc 780
tgccccacag catcatgat gggaaactcc ccttgggggt cgtctccccct tatgtgaaga 840
tgagttcggg gggctacacg gacccctga aattctacgc caccagctac tgcacagcct 900
acggtcggga ggatttcaag ccccggtgtg gcagtcacgt aggcaccggc tacaaatcaa 960
atttccagcc cgtggtctca tgccaagcca gtctggaggc cttagacaac ccggccaggg 1020
gggaacaagc ccaggaccat ttccagtctg tggccagcca gagctaccgc cccttgagg 1080
tgcttgacgg caagcatccc ctgcccctgga gcatgcgcca gaccagctca ggctatgggc 1140
gggagaagcc cagtgggggt cccccacca aggaggtccg gaaggtccat ttcgacaccc 1200
aggagcacgg gctcaggcc atcacggggc tggagcccag ggaagtgcct ctgtccacc 1260
agcagcaggg ccaggaccgg ctggagcggg agaacttccg acatggccca cgcttcatga 1320
cgtcggagta caattccaag tatctcaggg accctctaga ccagccagat ttcttgcaga 1380
agaaatcgat cggcgcaaag gaggggagtg gcttcaccaa acagtcccac cagagcccca 1440
ttgtcttcca accgcccctc caggcccctcc ctgggggaccc tggagatgag ttcctacctg 1500
tggtggccag aggttccaag cgggaaacag ccttcagccg agggaatgaa aggattctga 1560
accccagagt gccccctccc tgcccagaac ccagcagcgt gagtaccag cagtttcagc 1620
cacttcatcg gatgcaacag acaaacgttg ccttgcttgg ccgggagact gtggggaaaa 1680
aggagcccac agggttcagc cttaacaacc ccatgtatgt ccggagcccc tgtgatcctg 1740
acagggatca gcgatacctg accacctaca accaagggtg ctttgagaac atccccaagg 1800
gtctagacca ggaaggctgg actcgagggt gcatccagcc ccagatgcca ggaggctacg 1860
ccctcagcca gccggtcagc tgcattggag ccacccccaa ccccatggag agcctgcggc 1920
aactgcattc ccacgtggga agaaccctga cctcagctga ccccttctac cagaacacac 1980
ctcacagcag ccgctgcgtg gcacacagct gagccctgag cccttggttt gtcaatccta 2040
ggggacagga gccaggctac tgtctgtacc cccactccat ccccagggg aacaccctg 2100
cccagctact ttcctaataa aataaagagc attgaagcag agaa 2144

```

<210> 96

<211> 2117

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 4823849CB1

<400> 96

```

tggggctttc cctcatcccc aagttccaac caaggactgg gcaatgggtg cagatagttt 60
tacaaagtct agaacattgg ctgggcatgg tgggtcatgc ctgtgggtccc ggcacttttg 120
gaggttggga ggggaggatc acttgagccc aggagtttga caccagtctg ggcagcatag 180
tgagacccca tctctatttt ctaaaagggg tgtagagcat tgggagaagg ccctctggcc 240
ttgggaacat aggcgtccct cctcatatgt tcttatcca ggtccctgtg gctcctaagt 300
ggccacagct ctggtgcttt tctgccagcc gtgcacagtg aggtgggggg tgtcttagct 360
gagacacagt gccagtggt gggcatccag tctttctgtc ctgtgatggc cgccttctgc 420
ttcaagagaa taggtctcca tccactcaa gacccagcaa caagatagcc cctccatcat 480
cctgggagaa gtgtccaaa cctccctcaa ctcaagctgt ctgggtcttc cccattcaac 540
agctttttcc tacttttctt ctcttcttgg gactttcttc acacaatctc ccttttgggc 600
cactagaaac cagtgtctgg ttctcctgcc caaagatggg ggaatgaaaa cttcagtgcc 660
actcagcttt ttcttgctag tgcattttt tagttagata aaacagttat ctttttagtt 720
agataaaaac gaggaccac agaagaatga cacagggttc ctgcagccaa tgaccttaa 780
gtgtatttgg ggaggccaac gaaaaggatc aatagcaggt caggggccag ccagtggtg 840
gcacacctgc cccaagatg atgtgccact ttccagcatc tgtattccac cccttgttt 900
ctgtctcttt ctgacctatg tacctgaggc tgatttccag gtcactaagc ccgggaactg 960
gagagatgtc tgtgaagggt ctgccaccgt gatcctcggg gtcacctcct cgggtgccctc 1020
cctgccactc cccaatgtcc tctgatggc caatgttacc tggccccagg gtccatttac 1080
cacctggagc acactgggt acgccccagt catcaacctc agcagggttc tccccctgaa 1140
gtactgtggg ctacgaatct acgaccggct ccagcgcatc ctgaggggta ggacagtga 1200
cgaaaagatc tactatctga agctccacga aaaacaccca gagatttgtt ttcaattctg 1260
ggtcgcgttg gtgaaaattc tgcagaaagg cctgtccatc accatcaaag acccgagaat 1320
caaattcact cactgcctgg tgcccaagat gccaccaaac tccacagaaa caacacctga 1380
aaacagcctc ctgtcatccc cccagcccag cgagcccctc gtgctgctgg cggctgagca 1440
gaccagtggc agtttctcac agctctcagg aaagcccag ctcacagcag acaggaacaa 1500
tgacactgcc attgaaatag acaactgcag cagctacaag ataccctctc cagtggcatc 1560
tccaatcaat ttgaatatac ccatgagagc tgccttgagc cacagcctct gggaacaaga 1620
ggactggaat gagcaccttc tacaagttca tatagccagt tacctaggag agcacttctt 1680

```

```

gggagcctga cacctaggca aaacatgctc ttccttgcac acgtgtactt gctaccacct 1740
cttataatgc tccttttcca ctgtgagcca agtgaccttg ggagagagca acttttttgac 1800
acataaatct gacagaaagg accgggcacg gtgggtcata cctgtaatcc cggcactttg 1860
gaaggccggg gtgggcagat cacttgaagt cggaagtgtg agaccggcct ggccagcgtg 1920
gtgagacctt gtctctacta gaagtgtgaa aattagccag gtgtggtggt gcgcgtctgt 1980
ggccccggct gctcgggagg ctgaggcggg agaactctct gaactgggga ggccgagggt 2040
gcagtgagcc gagatcgctc cactgcgctc tagcctgggc aacagagcaa gattccatct 2100
caaaaaaaaa aaagggg 2117

```

<210> 97  
 <211> 1713  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 4433922CB1

```

<400> 97
cggttagctg aagaaagaca aaaagaagca gaaagaaaga agaaagaaga agaagaaaaa 60
tataacctgc aacttcagca ctactgtgaa agagacaatt tgattgggga agaaacaaag 120
cacatgagac agccttctcc tatagtccct gctcttcaga acaaaattgc aagcaaactc 180
caaagacctc cttcagttga cagcatcata cgctccttta ttcattgaaag ttccatgtcc 240
agggcacagt ccccccggt acctgccagg aaaaatcagc tccgtgcaga agaggagaaa 300
aaaaatgtaa ttatggaatt atcagaaatg agaaaacagc ttcgtagtga agagaggcgt 360
ctacaagagc gattgctaca catggacagt gatgatgaaa ttcctatcag gaaaaaggaa 420
aggaatccca tggatatatt tgatatggct agacatcggt tgcaagctcc tgtcagaaga 480
cagtccccta agggcttaga cgctgccact tttcagaatg ttcattgatt taatgagctg 540
aaagatagag attcagaaac acgagttgat ctgaaattta tgtacctgga tcctccaaga 600
gatcatcaca ccttagagat tcagcagcaa gccctgctaa gagagcagca gaagaggctg 660
aacagaataa aaatgcagga aggtgccaaa gttgacttag atgccatccc aagtgtctaa 720
gtacgagagc aaagaatggg gcttacgggt agacatatcc tgccattgaa gatgacgtcc 780
tccctccacc atcacagttg cctctgcac gggagcgcag gaggaacaaa tggaaaggac 840
tagacattga tagcagtcgt cctaattgtag caccagatgg tctctctcta aaatctatat 900
ccagtgtaaa tgttgatgag cttagagtga gaaatgagga acgaatgcga agactgaatg 960
aatttcacaa taaacctatt aatacagatg atgagagttc actggttgac cctgatgaca 1020
tcattgaaaca cataggggat gacggatcaa actctgtagc aactgagccc tggtccgcc 1080
ctggcacttc agaaacgctg aaacgtttca tggcagagca gctgaaccag gacgagcagc 1140
agattcctgg aaaaccaggc actttcactt ggcagggcct gtcgactgca catgggttaa 1200
ataaacctgt actggacca gtagtgccct ttaagggtgaa aggaatggta aatctgtacc 1260
tttaatatgt cctacttttg gccctacct gaaagttact ttttttccat catctgtata 1320
taaaattatt tttatcatga tgtatattat gtacataaat aaaaggccat gattattgat 1380
ttatataata gaattgtata gattatTTTT gcacagtttt gtcataaatt aggggtggtaa 1440
tgaactggat tgaactacta tatgtgcatt atattgaaat ctgcttggtca ttaagataag 1500
gtgaataagt gtcttaaacy tctgttaaaa ccggactccc ctttgttaca tgcacatttt 1560
ccattgttac ctcatgcaa aagaattcat ttagtaggta catctattgt agctgtgatt 1620
attocagttt ctgtgtgatg caatcaaatg tcctattaat tattatttca tgtcatttgt 1680
agctactgat acagcagaaa tgaagggaac tgt 1713

```

<210> 98  
 <211> 1297  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7504597CB1

```

<400> 98
cacgcgcgca cgggatcctg ctccgctccg cggctcccc gccccggaga gctccgcgat 60
cagcaccggg gacagcgcca ccgcccacgt gcaggggggt ggggtccggc ggggctggcg 120
cctcggcgct tcccgggagt gtcccgctcca gccgcgcgag aggggtgctt agtgtctctg 180
cagaaaagac tccaggacct cgccaccatg tccccggagc cccaacccc ggggcctcca 240

```

```

tcgcccagaca cgccctcccga ctccagtcgc atcagccacg gccaggtgc agccagaagt 300
agaggagctg gagccagcac catccgggccc agggcagcag gtggcagaca agcatgagct 360
aggacgactg cagtactccc tggattatga ctccagagt ggccagctgc tgggtggcat 420
tctgcaagca atgggattgg cagccttgga tcttggtggc tcctcggacc cctatgtgcg 480
ggtctacctg ctgccggaca aacggaggcg gtacgagacc aaggtgcatc ggcagacgct 540
gaaccctcac tttggggaga ccttcgcctt caaggtcccc tacgtggagc tggggggcag 600
ggtgctgggtc atggcgggtg acgacttcga ccgcttctct cgcaatgacg ccatcgggga 660
ggtgcggggtc cctatgagct ccgtggacct ggggcggcca gtgcaggcct ggcgggagct 720
gcaggcggtc ccgcgggagg agcaggagaa gcttggggac atctgcttct ccctccgcta 780
tgtccccacg gccgggaagc tcaccgtcat cgtcctggag gctaaaaacc tgaagaagat 840
ggacgtagga ggactgtcag atccatacgt caaggtccac ctgctgcagg gcggaacaaa 900
ggtgcggaag aagaaaacca ccatcaagaa gaacactctg aaccctattt acaacgaagc 960
tttcagcttc gaggtgccct gtgaccaagt ccagaagggtg caggtggagc tgaccgtgct 1020
ggactacgac aagctggggc agaacgaggc catcgggagg gtggccgtgg gggcgggcgc 1080
cggcggggct ggcctgcggc actgggcgga catgctggcc aaccgcggc ggcccatgtc 1140
ccagtggcac tcgtgcggc ccccggaacc agtgaggctg ctgcctgcgc cctgactccc 1200
accccaagcc cccggccaag cctggactct agccctgac ccagacttc cgagaccggc 1260
caaagccaag ccagccccac ccttaagctt cctctga 1297

```

<210> 99  
 <211> 885  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7505987CB1

```

<400> 99
gccaggagg cgccccgggt gaaggcacgg gtgcgcacga ggagttccgg ctggagaccc 60
gtgctctggg ccggcgccct caccatggcc tcggcagagc tggactacac catcgagatc 120
ccgatcagc cctgctggag ccagaagaac agccccagcc caggtgggaa ggaggcagaa 180
actcggcagc ctgtggtgat tctcttgggc tggggtggct gcaaggacaa gaaccttgcc 240
aagtacagt ccactacca caaaaggaa ctgctcgagc tgctctttga ttatgagatt 300
gagaaggagc cctgctctt ccatgtcttc agcaacggtg gcgtcatgct gtaccgctac 360
gtgctggagc tcctgcagac ccgtgccttc tgccgcctgc gtgtggtggg caccatcttt 420
gacagcgctc ctggtgacag caacctggta ggggctctgc gggccctggc agccatcctc 480
gagcgccggg ccgccatgct gcgcctgttg ctgctgggtg cctttgccct ggtggtcgtc 540
ctgttccacg tcctgcttgc tcccatcaca gccctcttcc acaccactt ctatgacagg 600
ctacaggacg cgggctctcg ctggcccagag ctctacctct actcgagggc tgacgaagta 660
gtcctggcca gagacataga acgcatgggt gaggcacgcc tggcacgccg ggtcctggcg 720
cgttctgtgg atttctgtgc atctgcacac gtcagccacc tccgtgacta ccctacttac 780
tacacaagcc tctgtgtcga cttcatgcgc aactgcgtcc gctgctgagg ccattgctcc 840
atctcacctc tgctccagaa ataaatgcct gacacctccc cacaa 885

```

<210> 100  
 <211> 1652  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7506025CB1

```

<400> 100
cagcacttcc ggtacggaaa actcgtgct gccccaacct ggcttgacag gcttggctctc 60
tgcaagtggc tctcagcccc ttcttctttc ctgcctcacc ttccaattcg tttgccgccc 120
ccgtcccgcg gctgctgttt ccggagttgc cccttcccca tgttccgggg caggagtcgg 180
caaagcgaag atccgcccgc cggttcccca tcatgtccga actgactaaa gagctgatgg 240
agctggtgtg gggcaccaag agcagccccg gtctctcgga caccattttc tgccgctgga 300
cgcaagggtt tgtgtttagt gaatcagagg gatctgcatt agaacagttt gaaggtggcc 360
cctgtgctgt tattgcacct gttcaggcat ttcttttgaa gaagctcctg ttttcttcgg 420
agaagtcttc ttggcgggat tggtcagagg aagagcagaa ggaactcctt tgtcatacct 480

```

```

tgtgtgatat tttagaaagt gcttgttgtg accactcttg atcatactgc ttggtttcat 540
ggttaagagg aaagacaact gaggaactg ctagtatctt tgggagtcct gcagagtccta 600
gttgccaagt ggaacattct tctgccttgg ctgtcgaaga gcttggcttt gagcgatttc 660
atgcattaat tcaaaaaaga tctgttcagaa gtttaccaga attaaaagat gctgtcttgg 720
accagtattc aatgtgggga aataaatttg gagtattgct ttttctgtat tctgtattac 780
tgacaaaggg cattgaaaac ataaaaaacg aaattgaaga tgcaagtga cccttgatag 840
atcctgtata tggacatggc agccaaagt taattaatct cctgctgacg ggacatgctg 900
tttctaagt atgggatggg gatagagagt gctcaggaat gaaacttctt ggtatacatg 960
aacaagcagc agtaggattt ttaacactaa tgggaagctt aagatactgt aaggatatgg 1020
ctttagttgc ccctgaagct ccttcagaac aagccagaag agtttttcaa acctacgacc 1080
cagaagataa tggattcata cccgattcac ttctggaaga tgtgatgaaa gcattggacc 1140
ttgtttcaga tcctgaatat ataaatctca tgaagaataa attagatcca gaaggattag 1200
gaatcatatt attgggcca tttcttcaag aattttttcc tgatcagggc tccagtggtc 1260
cagaatcttt tactgtctac cactacaatg gattgaagca gtcaaattat aatgaaaagg 1320
tcatgtactg agaagggact gcagttgtga tgggttttga agatcccatg ctacagacag 1380
atgacactcc tattaacgc tgtctgcaaa ccaaatggcc atacattgag ttactctgga 1440
ccacagatcg ctctccttca ctaaattaat ttgtctaagt atttataagg aagatcttaa 1500
taacagatgt tgaaagaagg agtcaagact ggcaattggc tggattaagc taaacactgg 1560
tatcactgat taactgtaaa taacaattaa aaacacattt tcagtgttta tgatatgttt 1620
aaattatttg tcctaaagct ttatgtttaa ga 1652

```

<210> 101

<211> 1022

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7506102CB1

<400> 101

```

cagccagctg tcccagagtct ggactttccc tctgccctc cccactctca ggctgggtggg 60
gtggggaaaag cagcccattc ctgggctcag agactccac cccagctcag agggagcagg 120
ggcccagcca gggacggacc ctcatctctc ccagggacct cagacctctg tctctctcgg 180
gtaagtctcc atctctgtct gtctctgtct ctgtctctgt ctctgtctgt ttttcacgca 240
ctcagcaagg cctcctgccc tgagagaggc tccgcccact acccccact ttcccataa 300
aaccagctga gtatttgtgc caggaagact gcgtgcagaa ggtgactgtc tcagtggagc 360
tgggtcatct cagcaggcct tggctccttg aacttttggc cgccatgtgc ttcccgaagg 420
tcctctctga tgacatgaag aagctgaagg cccgaatgga gctcactcct ctacttgaaa 480
aagaaagaga tggattacgg tgccgaggca acagatcccc tgtcccggat gttgaggatc 540
ccgcaaccga ggagcctggg gagagctttt gtgacaaggt catgagatgg ttccaggcca 600
tgctgcagcg gctgcagacc tgggtggcac gggttctggc ctgggtgaag gagaaggtgg 660
tggccctggg ccatgcagtg caggccctct ggaaacagtt ccagagtttc tgctgctctc 720
tgtcagagct cttcatgtcc tctttccagt cctacggagc cccacggggg gacaaggagg 780
agctgacacc ccagaagtgc tctgaacccc aatcctcaaa atgaagatac tgacaccacc 840
tttgccctcc cgtcaccgc gcaccacccc tgaccctcc ctcagctgtc ctgtgccccg 900
ccctctcccc cacactcagt cccctgcct ggcttctctg ccgcagctct gacctgggtc 960
tgtcgccctg gcattctaat aaaacctgct tatacttccc tggcagggga gataccatga 1020
aa 1022

```

<210> 102

<211> 2046

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 1333949CB1

<400> 102

```

gccagttaag aactcacagg tactttttcca gccaccaga taggagagat cattaaaaca 60
gtgcattctg tgctacctga cacctattgg ggtcctggaa ggaggaagca acaatcctga 120
gtgaaacctc gacaagaagt atccaatagg acattcgtca tgtcctcaaa tacaatgctt 180

```

caaaaaacgc	tgctgatctt	gatctctttt	tcagtagtaa	cctggatgat	ttttataatt	240
tctcagaact	tcacaaagct	ttgggtctgt	ctaaacttat	ccatctctgt	ccattactgg	300
aacaactccg	caaagtcctt	attccctaaa	acatcactga	taccaattaa	gccactaaca	360
gagactgaac	tcagaataaaa	ggaaatcata	gagaaactag	atcagcagat	cccacccaga	420
cctttcacc	atgtgaacac	caccaccagt	gccacacaca	gcacagccac	catcctcaac	480
cctcgagata	catactgcag	gggagaccag	ctggacatcc	tactggaggt	gagggaccac	540
ttggggacaga	ggaagcaata	tgggtggggat	ttcctgaggg	ccaggatgtc	ctccccagca	600
ctgacggcag	gtgcttcagg	aaaggtgatg	gacttcaaca	atggcaccta	cctgggtcagc	660
ttcactctgt	tctggggagg	ccaggtctcc	ctgtctctgc	tgctcatcca	ccccagtga	720
ggggcgctcg	ctctctggag	ggcaaggaa	caaggctatg	ataaaaattat	tttcaaaggc	780
aaatttggtt	atggcacctc	tcatgtcttc	actgaatgtg	gcctgaccct	aaactcaaat	840
gctgaactct	gtgaatatct	ggatgacaga	gaccaagaag	ccttctattg	tatgaagcct	900
caacacatgc	cctgtgaggc	tctgacctac	atgaccaccc	ggaatagaga	ggtatcttat	960
cttacagaca	aggaaaacag	ccttttccac	aggatggcac	ctggagagac	atctatagca	1020
ggaaatcagg	tccaaagtgg	gagttgaaat	gatgaaggat	cgtaaacaca	ttgatgtcac	1080
taattgtaac	aagagagaaa	aaatagaaga	gacatgccaa	gttggaatga	agcctcctgt	1140
ccttggtggt	tatactttac	aaggaaaatg	gataacaaca	ttttgcaacc	aggttcagtt	1200
agacacaatt	aagataaatg	gctgtttgaa	aggcaaaactc	atttacctcc	tgggagactc	1260
tacactacgt	cagtggatct	actacttccc	caaagttgta	aaaacactga	agttttttga	1320
tcttcatgaa	actggaatct	ttaagaaaca	tttgcttctg	gatgcagaaa	gacacactca	1380
gattcaatgg	aaaaaacata	gctatccctt	cgtcactttc	cagctctact	ctctgataga	1440
tcatgattat	atccctcggt	aaattgaccg	gctatcaggt	gacaaaaaca	cagccatcgt	1500
catcaccttt	ggccagcact	ttagaccatt	tcccattgac	atttttattc	gcagggccat	1560
cgggtgttcaa	aaggctattg	aaagactggt	cctaagaagc	ccagccacta	aagtgtattat	1620
taagacagaa	aacatcaggg	agatgcacat	agagacagag	aggtttggag	acttccatgg	1680
ttatattcac	tatcttatca	tgaaggatat	tttcaaagac	ctcaacgtgg	gcatacttga	1740
tgcttgggac	atgaccattg	catatggcac	tgacactatc	cacccacctg	atcatgtgat	1800
tggaatcag	attaacatgt	tcttaaaacta	catttgctaa	gggataaata	ctatacaaaa	1860
tcactaggaa	ccaatctctg	cacataatcc	cacatgtatt	gtaaaagtaag	ttttactcat	1920
tttaggaact	aaggaaaata	aattttaaag	aatctgtttg	gggaggaagg	ctatgtaagg	1980
acaatgacaa	ctgataaggg	atgcaaaacc	aagagaatca	ttcaatgaag	aagtggacta	2040
tacca						2046

&lt;210&gt; 103

&lt;211&gt; 2355

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7035533CB1

&lt;400&gt; 103

ggtgaatata	accactgata	gtcgcattgcc	gcgtacgtca	gctcggtaat	acggctcgag	60
tgatgatgaa	gtgatttgca	atagtatggg	ggagcatctc	ttcagaatgg	gcttttagga	120
tgaagcagct	gggtattttc	ataaagcagt	gaagctaaac	cctgatttca	gtgatgcaaa	180
ggagaatttt	tatcgtgttg	caaactgggt	ggtggaacgc	tggcacttta	tcatgcttaa	240
tgacaccaag	aggaatacaa	tttataatgc	agcaatccaa	aaggcagttt	gtttgggggc	300
caaaagtgtt	ttggacattg	gagcaggaac	tggaatacta	agcatgtttg	ctaaaaaagc	360
tggagcacat	tccgtgtatg	cctgtgagtt	atccaagacc	atgtatgaac	ttgcctgtga	420
tgctgtggca	gcaaacaaga	tgggaagcag	gatcaaaactc	ttacatacga	agtcacttga	480
catagagatt	ccaaaacata	ttcccgaag	accaaaagggt	aaagtgtctaa	ttgtgaaaag	540
tatgggaaaag	ttataaccagc	aagtgtctgt	atattttggga	tggcagtaga	atgtgcagag	600
ataagaagac	atcatagagt	gggtattaa	gacattgtctg	gtatccattt	gccaacaaat	660
gtgaaatttc	agagtcgggc	ttattcttct	gtagatactg	aagaaacaat	tgaaccttat	720
acaactgaaa	agatgagtcg	agttcctgga	ggatattttg	ctttgacaga	gtgctttgaa	780
attatgacag	tagatttcaa	caaccttcag	gaattaaaaa	gtcttgcaac	taaaaagcct	840
gataagattg	gtattcctgt	tattaaagaa	ggcatactag	atgctattat	ggtttggttt	900
gtgctccagc	ttgatgatga	acatagttta	tccacaagtc	ctagttagga	aacatgttgg	960
gaacaggctg	tctacccgt	acaggacctt	gcagactact	ggataaagcc	tggagaccat	1020
gtgatgatgg	aagtatcttg	tcaagactgt	tacttaagaa	tccagagtat	tagtgtcttg	1080
ggtttggaat	gtgaaatgga	tgttgcaaaa	agttttaccc	agaataaaga	cttggttatcg	1140
ttaggaaatg	aggctgaact	ttgtagtggc	ctcgctaacc	ttcagaccag	taaaccagat	1200

```

gctgtagagc agacatgtat attggaatct acagaaattg ctttgcttaa caacatccca 1260
tatcatgaag gctttaaaat ggcaatgagc aaagttttgt cttcactgac tccagagaaa 1320
ctgtatcaga ccatggatac tcactgtcag aatgagatga gctctggaac tggacagagt 1380
aatactgtac agaaccatcct tgaacctttc tacgtgttag atgtgtccga aggcttctct 1440
gttctgcctg ttattgctgg cacacttggg cagggttaaac catacagttc tgtggagaaa 1500
gaccagcatc gtattgctct ggacctcata tctgaagcca atcactttcc taaagaaaaca 1560
cttgagtttt ggctgagaca tgtggaggat gaatctgcta tgttacaaag gccaaaatca 1620
gacaagttat ggagcataat tatattggat gtcattgagc catctgggct cattcagcag 1680
gaaataatgg aaaaagctgc aatatccagg tgtttactac aatctggagg caagatcttt 1740
cctcagtatg tgctgatggt tgggttgctt gtggaatcac agacactcct agaggagaat 1800
gctgttcaag gaacagaacg tactcttggg ttaaataatag caccttttat taaccagttt 1860
cagggtaccta tacgtgtatt tttggacctt tctcattgct cctgtatacc ttaagcaag 1920
ccagtggaa ccttaagact agatttaatg actccgtatt tgaacacctc taacagagaa 1980
gtaaaggat acgtttgtaa atctggaaga ctgactgcta ttccattttg gtatcatatg 2040
taccttgatg aagagattag gttggatact tcaagtgaag cctcccactg gaaacaagct 2100
gcagttgttt tagataatcc catccagggt gaaatgggag aggaacttgt actcagcatt 2160
cagcatcaca aaagcaatgt cagcatcaca gtaaagcaat gaagagcagt tttccaatga 2220
aaactgtgta aatagagcat caacaagtac aaaattcttg tcttaattag tgggggtata 2280
taaaaattcc ttgtaatggg caaatatatt ttaaaattga cathtaataa gcataatttta 2340
aaagataaaa aaaaaa 2355

```

&lt;210&gt; 104

&lt;211&gt; 2573

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2815375CB1

&lt;400&gt; 104

```

cggctcgaggc ggcgggcgcg ggacgggttg tccagttctc cggcctggcg gcaggcaagt 60
ctagctcggc gctgtcggat acttgggggt agcggaaaagc atggcgggga cctccgcgcc 120
aggcagcaag aggcggagcg agcccccg gcctcgcccc ggcccgccgc cgggcaccgg 180
gcaccccccg agcaagcggg cccggggctt ctccgcagcc gctgccccgg accctgacga 240
cccgttcggc gcgcatgggg acttcaactg cgacgacctg gaggagcttg acaccctcgc 300
gtcacaggcc ctgagccaat gtccggccgc ggctcgggac gtgtccagt atcataaggt 360
ccacagatta ttagatggca tgtcaaaaaa tccctcaggg aaaaacagag aaactgttcc 420
aattaaagat aatttcgaat tagaggtagt tcaggcacaa taaaaagaac ttaaagaaaa 480
gatgaaagta atggaagaag aagttctcat taagaatgga gaaattaaaa ttttgcgaga 540
ctcactacat cagacggaat ccgttctaga ggaacagaga agatcacatt ttcttcttga 600
gcaagagaaa acccaagcac tcagtgcaca ggaaaaggaa ttctccaaaa agctccaatc 660
attgcagtct gaactccagt ttaaagatgc agagatgaat gaattaaaga caaagctcca 720
gaccagtga cagacaaata aactggctgc tccctctgtt tcccatgtca gtccatagaa 780
aaacccttct gtggttataa agccagaagc atgttctcca caatttggaa aaacatcttt 840
tctacaaaag gagtctttta gtgctaacat gtcccttccc caccctgcc agacggagtc 900
aggatacaag cctctgggtg gcagagagga tagtaagccc cacagtctga gaggtgactc 960
cataaaacaa gaagaggccc agaaaagctt gtgtgacagc tggagacaga gatcaaacac 1020
tcaaggttcc attttgataa acctgctcct gaagcagcct ttgatcccag ggtcatccct 1080
aagcctttgc cacctcctga gtagtagttc tgagtctcct gctggcacc cctgcagcc 1140
accagggttt ggcagtacct tggctggaat gtcaggcctc aggaccacag gttcttatga 1200
tggtgcatct tccctctcag ccctgagaga agcacagaac ctggcattca ctggactgaa 1260
tctggttgcc cggaatgagt gctcacgtga tggagaccca gcagagggag gcagaagggc 1320
cttcccactc tgccagcttt ctggagccgt gcatttctc ccccttgtag agttcttcat 1380
cggcttacac tggcaggccc tgcaggactt ggcagctgct aagagaagcg gagcacctgg 1440
ggactcaccg acacattcct cctgcgtgag ctctggggta gagaccaacc ctgaggactc 1500
agtgtgcatc ctggaaggct tctctgtgac tgcacttagc attcttcagc acctgggtgtg 1560
ccacagcgga gcagtcgtct ccctattact gtcaggagtg ggggcagatt ctgctgctgtg 1620
ggaaggaaac aggagcctgg ttcacaggct tagtgatgga gatatgacct cagccctaag 1680
gggggttgct gatgaccaag gacagcacc ctggtgaag atgttcttc acctgttggc 1740
tttctcttct gcagcaacag gtcaccttca agcagtgct ctgacccagt gccttaaggt 1800
tttggtgaaa ttagccgaaa acacttcctg tgatttcttg cccagggttc agtgtgtgtt 1860
ccaagtgtct ccaaagtgcc tcagcccaga gacaccctg cctagcgtgc tgctggctgt 1920

```

```

tgagctcctc tccctgctgg cggaccacga ccagctggca cctcagctct gttcccactc 1980
agaaggctgc ctcctgctgc tgctgtacat gtacatcaca tcacggcctg acagagtggc 2040
cttgagagaca caatggctcc agctggaaca agaggtgggtg tggtccttgg ctaagcttgg 2100
tgtgcagagc cccttgcccc cagtcactgg ctccaactgc cagtgtaatg tggaggtggg 2160
cagagcgctc acggtgatgt tgcacagaca gtggctgaca gtgcggaggg cagggggacc 2220
cccaaggacc gaccagcaga ggcggacagt gcgctgtctg cgggacacgg tgctgtgct 2280
gcacggccta tcgcagaagg acaagctctt catgatgcac tgcgtggagg tcctgcatca 2340
gtttgaccag gtgatgccgg gggtcagcat gctcatccga gggcttccctg atgtgacgga 2400
ctgtgaagag gcagccctgg atgacctctg tgccgcggaa accgatgtgg aagaccccga 2460
ggtggagtgt ggctgaggcc ctgagtgtcc agccacatgg tggcaccagc accactcctt 2520
tccttaccac atcaactgat taaagcagtg accagcagga actgcccaga aaa 2573

```

&lt;210&gt; 105

&lt;211&gt; 1488

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2820152CB1

&lt;400&gt; 105

```

ccccggggga ctggggcgggc ggaggccggg ctgccgaagt gtgagggggag gagcccgagc 60
cgctcgctgct gccgcgcgcta agggagagga gcatgtctgc aactcgagcc aagaaagtga 120
agatggccac caaatcatgc ccgagtgctg accaacaggt tcctgttgca tgtaaatcat 180
gtccttgttg ttacatatatt attagcagaa aactttttaa tgcaaaacac tcagagaaat 240
caccaccttc tacagaaaac aagcatgagg ccaagaggag gcgaacagag agagttagga 300
gagagaagat aaattctaca gtaataaaag atttagaaaa cagaaagagg tctcgaagta 360
acagccattc agatcatatc agacgaggaa gaggaagacc taaaagtgc tctgccaaaa 420
aacatgagga agaaagagag aaacaggaaa agggaaattga catctatgct aacctgtctg 480
atgaaaaggc tttcgtgttt tcagtcgcct tggcagaaat aaatagaaaa attatcaatc 540
aaagacttat tctctgatac ttgtttgcaa aatctgtctg aacttctgtc aggataacat 600
cagcaaattc tcaaacagtt atggactctc aggagcattc tgactgcac aataagggcc 660
atcgattgtt tctttctcgc ctgtcgttta gcctgtgcc cacttcggga gcaaagctgt 720
gtaggcttgg actgttcttg gatttccttg tttaccaagg agtattgtca gtgttttgtg 780
tttgggatat aagtaagtgt atttgccaac aacaaaaaaa aaccaacaaa aatgctggat 840
ggatgttaaa ggataaaaag tgatggggta tggggtaggt atgaaaacta gggtaataag 900
atgcttgatt gttttcaggt ttgccaccag agatgcaata ttttaaatac tgtcagaaag 960
agtgtttttt aaaaatatat atttcagatt ttcagtgcata acagattgca tatatacagt 1020
tcatttactt ttattaaatt ggcaatcgat attttattga atggggaatt aagaatataa 1080
actgtaaatt ttgtcttaaa ttgatttttt ttcttggcct ttatttttat tacttaaggt 1140
atgagttata ttttaggaga tactttctag acagtttgaa gtttgggtgc tatatgaaga 1200
aaaacattta atgatatact cagatataaa caccttgttt tcttgaatgt aatttattta 1260
ttagctaaaa aaaaaaaca aactgtcttc aagatgggtg ctgttaacct aaacatgaaa 1320
tatattttat acatgcaaaa tatttcttag ccacttttaa gcagtaattg tttatctgaa 1380
tggaattttg tttctatata ctgataatta tttaaattgt tactttcaat ctggaagaaa 1440
aaataaacct taatattgaa actaaaaaaa aaaaaaaaaa aaaaattg 1488

```

&lt;210&gt; 106

&lt;211&gt; 4486

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2959305CB1

&lt;400&gt; 106

```

cggtagctgtg gcagcagtc gtgtgtctgg cgggtgtcgg cgtgaagcgg ggctggggcca 60
gcgggaggga gctctgtggg agtggaaggc ctgtatttct ctacctggac cgcacgctcc 120
tcgcgcaagg agggtagagc tcaagatgct tgggatttgc ttacattct aagtggattt 180
ggagaagaga caataaagaa taagcagtc attctataat atatttagaa ggagaaaagt 240
gccaaaggagg aatgaaaaag tagagcaaga gacagcatct tgctgccacc caggctggag 300

```



tgacgtgggtg	cagtcttagc	ttactgcagc	ctcaaacttc	tggggtcaag	taatcctccc	360
gcctcagctt	cctgagtagc	tgagattaca	ggttttatgg	caattggata	tattccgacg	420
aagcttgccg	gttttgactg	gacatgtttg	tcaggagat	gcctgtatat	tttgtgcatt	480
gaagacgata	tttgcacagt	tccaacacag	tcgagaaaaa	gcacttccct	cagataacat	540
aaggcatgct	cttgcagaaa	gtttcaaaga	tgagcagcga	tttcaacttg	gccttatgga	600
tgatgctgcg	gagtgctttg	aaaatatgtt	ggagaggatt	cattttcaca	tagtgccaag	660
cagagatgca	gacatgtgta	cctctaaatc	ttgtatcact	caccagaagt	ttgctatgac	720
tctgtatgaa	cagtgtgtgt	gtcgtagctg	tggagcatcg	tcagatcctc	taccttttac	780
agaatttgtg	cggtagatct	ctacaacagc	cttatgcaat	gaggttgaaa	gaatgttgga	840
aaggcatgaa	cgctttaaac	ctgaaatgtt	tgcagaattg	ctacaagcag	caaatacaac	900
agatgactat	aggaaatgtc	ctagtaactg	tggccaaaaa	ataaaaattc	gccgtgtttt	960
aatgaattgc	ccagagattg	ttacgattgg	tttagtctgg	gactccgagc	attctgactt	1020
gaccgaagct	gttggttcgga	atctagcaac	acatctttat	cttctctggg	ttttttatag	1080
agttactgat	gaaaatgcca	aaaatagtga	acttaacctt	gttggtatga	tctgctacac	1140
cagccaacat	tattgtgcct	ttgcatttca	caccaaaagt	tccaaatggg	tattttttga	1200
tgatgcaaat	gtgaaagaga	ttggaactag	atggaaagat	gttgtctcca	aatgcattcg	1260
atgccacttt	cagccactac	ttttgtttta	tgcaaaccca	gatggcacag	cagtttctac	1320
tgaggatgca	ctcaggcagg	tcacagctg	gtcacattac	aaatctgttg	cagaaaaatat	1380
gggatgtgaa	aagcctgtaa	ttcataagtc	agataattta	aaagaaaatg	gattttggtga	1440
tcaggccaaag	cagagagaaa	atcagaaatt	tccaactgat	aataatttcat	catctaactg	1500
gagccacagt	cacacaggtg	tagggaaagg	accagctaag	ttaagtcaca	ttgatcaaaag	1560
ggaaaagata	aaagacattt	ccagagaatg	tgctctgaaa	gctattgaac	agaaaaactt	1620
actttcttca	caaaggaaag	atthagagaa	gggacaaaga	aaagatttag	gacgacatag	1680
agattttggtt	gatgaagacc	tttcacattt	ccaatctgga	tcacctcctg	ccccaaatgg	1740
ttttaaacaa	catgggaatc	cacatctata	tcatagtcaa	ggaaaaggat	catataaaca	1800
tgaccgagtt	gtacctcaga	gtcgagcttc	tgacaaaata	ataagttcaa	gtaaatccca	1860
gattcttgc	ccaggagaga	aaataactgg	caaagttaag	agtgacaatg	gcactggata	1920
tgacacagac	agcagccaag	attctagggg	tagaggaaac	agctgtgata	gcagcagtaa	1980
aagccggaac	cgaggttgga	aacctatgag	agaaacatta	aatgttgata	gtatttttag	2040
tgaaagtgaa	aaaagacagc	atagtccaag	acataaaacca	aatatcagta	ataagcctaa	2100
atctagcaag	gatccgagtt	ttagtaattg	gccaaaagag	aatccaaagc	aaaaagggtt	2160
aatgaccata	tatgaagatg	aaatgaagca	ggaaatagga	agcagaagtt	cccttgaatc	2220
taatggaaaa	ggagcagaga	aaaataaagg	ccttgtagag	ggtaaagtgc	atggtgataa	2280
ttggcagatg	caaaggactg	agtcctggata	tgaaagcagt	gatcacatca	gtaatgggtc	2340
tactaattttg	gactcacctg	ttatcgatgg	aaatggtaca	gtaatggata	tcagtgggtg	2400
taaagaaaaca	gtatgcttca	gtgaccagat	tacgacaagc	aacctaaata	aagaacgtgg	2460
ggactgtacc	tccttccaga	gccaacatca	cttagaaggc	tttagaaaag	aactcaggaa	2520
tttggaagca	ggctataaat	ctcatgaatt	ccaccagaa	tcacatttac	aaataaaaaa	2580
tcattttgata	aaaagatcac	atgtacatga	agacaatgga	aaagttatttc	cttcatccag	2640
tctacaaaata	ccaaggacc	ataatgcaag	agaacatatc	caccagtcag	atgaacagaa	2700
acttgaaaaa	ccgaatgaat	gcaaattttc	tgagtggctt	aatatagaaa	attctgagag	2760
aacaggtttg	ccttttccag	ttgataactc	tgcttctggg	aagagagtga	acagtaatga	2820
accatcttca	ttatggtctt	cacacctaa	aactgttggg	ttaaagccag	aaactgtctc	2880
tctcatccag	caacaaaata	tcattggatca	atggttacttt	gagaactctc	tatccacaga	2940
atgtataatt	cggctcagcca	gcagatctga	tgggtgtcag	atgccaaaac	ttttttgcca	3000
gaatctacca	cccccttgc	caccaaagaa	atatgtctata	accagtgtgc	cacagtcaga	3060
gaaaagcgaa	tctacacctg	atgtcaaaact	tacagagggtg	tttaaagcta	cctctcatct	3120
tcogaagcac	agtttaagta	cagcttcaga	accaagttta	gaagtgaagta	catataagaa	3180
tgatgaaaga	cataaagaaa	catttcaagt	gagagaatgt	tttggcaaca	caccaaactg	3240
tccatccagc	tcctcaacta	acgatttttc	ggcaaaactca	ggtgccattg	atgcattttg	3300
ccaaccagaa	ctagactcta	tttctacctg	tccaaatgag	acagtttcat	taactacctc	3360
tttttcagtt	gatagctgca	tgacggatag	atatagattg	aaataccatc	agaggcccaa	3420
gctctctttt	ccagagagca	gtggcttttg	taataattca	ctatcttaga	gtgaaaaagg	3480
actagacctg	tggtacataa	taactcttgt	tcaagtgc	cttctgaaca	aagatataaa	3540
cctagcatal	attgtaatat	ataactggta	aaactgacca	acttttactt	ctcagaggcc	3600
atttaaatat	aataggaacc	tactgaccaa	acctagtgat	acataaaaatt	aaagcctgtg	3660
gcatttttaa	agttgttaat	cactatacat	atgtatgtgt	atatgtgtat	acacatatat	3720
aatttttatga	tcaatatctt	agatatattta	gaaattccct	ttgaatagtc	ttggcggtgc	3780
gtgaaaaata	gaaaaatcag	gagatataat	aattcatttg	tcatatgcta	cagttgaata	3840
aaaattaaat	ttgtccccct	acttagactg	cagtagactg	ggaagtatta	aactaaccag	3900
tactcatgta	tggagattat	aaactatgct	aattgaaaag	taccaggtga	tacatagcct	3960
ggtacttttt	tgtcaaagaa	ttgctttatg	aagaagcact	ttctaattgg	taaccagaaa	4020
agggtctcac	ccatttgttt	ttgacttcta	tgatagtcac	tgcatatgat	ccctttaagt	4080

```

gtcttttaaaa aaaatgactt atgaatttga tagcatttgg ggaaaaaaag ccttatatga 4140
ttaacaattt gtaagttttc aagatcggat ggaattttaat tttgctccta gaatttttgt 4200
cttcagaata tctagttaag ataaattagg cctttgacta ttataggtat tcataaatgc 4260
tacttttagcc atgtaagtta aaaagttaaa atacttaaac tttagtgtat ttattttatct 4320
cgttactttt tttctaattt tatatgaaat gtgaaaaggtc tttatttttg tttgggtttac 4380
tttgggctgc taaccaccag tgtaggaat taatgtggg caccctctat tgggaatgtc 4440
actgtaattg cctacatttt ctttgctgta atattacctg ctaata 4486

```

<210> 107  
 <211> 1223  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 4913449CB1

```

<400> 107
tgagaccatc cccatcccca gggatgtata taagggtgact ttgctccatc ctgggtattt 60
tgatgagtag tccaaaggct gagctgaaac tggagaagaa gtcagggcgc aagccaagga 120
gccccaggga cagtggggcca cagaaagaac tgggtgatccc agggattgtg gatttcgagc 180
ggatccgccc ggcatgtagg accccaaagc cccaaacccc tggcacctac tgctttggac 240
gcctcagtca ccactccttc ttctcccggc accaccacac cccccagcac gtgaccacac 300
tccaagatct caccgggaag cctgtctgtg tcgtcaggga ttttccagcc cccttgcttg 360
agtcaactgt cttttccggc tgtcaaattg ggatacccac catctctgtc cccattggag 420
acccacagtc taatcggaac cccagccttt cttctgaagc ctggaagaag gagttgaagg 480
agctagcttc ccgggtggcc ttctcacca aggaggatga actgaagaag aaagagaagg 540
agcagaagga ggagcctctg cgggagcagg gggcaaagta ctacagagag actgggagagc 600
tcatccccgc ttccaccggg gctgtcggcc gccgcagatc tcaccagggc cagcagagtc 660
agtcttccag cagacatgaa ggagtccagg ctttctctct tcaggatcag gagctgctgg 720
tcctggagct cctgtgtcgg atcctggaaa cagacttgct aagcgcaatc cagttctggc 780
tgctctacgc tccgcccacg gaaaaagacc tcgctctggg actcctgcag acagcagtg 840
ctcagctcct tccccagccc ctagtctcca tccctacaga aaagctccta agccagcttc 900
cagaagtgca tgaacctcct caagagaagc aagagccacc ctgcagtcaa tccccgaaga 960
aaacgaagat atcacctttt acaaaaagcg aaaaaccaga gtacattgga gaagctcaag 1020
tcctccagat gcattcaagc cagaacacag agaagaagac atcgaagccg agggcagaga 1080
gctgaggggc cctaacactt gcacctgcct tgctcaagag cagccccaa ggttcagggg 1140
tgtttctgtc tccaccactc tcacagcagt acctgattcc ctaccgtgaa aactcttact 1200
aaataaaaacc gtcttccctg aga 1223

```

<210> 108  
 <211> 1188  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7506136CB1

```

<400> 108
ggagactttc caggctgggc tggacacacc tttccaagga cccccaaact ctgctccgtg 60
cacgtcaaat gtccttttcc cttgtgtcca accccctacc cctctcccta acaccctct 120
tctcaacaag actcagcttc tccccgaggt gggtagcat ccttgaggtt tcccaccctt 180
aactgctgtg tccccggatg gagccagaga aatgtgggtg gggggccggg gccagagttt 240
caacattgcc cccagaagg aggagccaga gatggggtct gtccaggaaa acaggatgcc 300
ggagcccagg agtcgtcagc ctagcagttg cctggcctcc agatgcctcc caggggagca 360
gatcctagca tgggccccag gggtaggaa gggcctggaa ccagaattgt ctggaaccct 420
gatctgtacc aactttaggg tcaccttcca gccctgtgga tggcagtgga atcaggacac 480
tcccttgaac agtgaatacg attttgccct ggtcaacatt ggacgattag aggctgtgag 540
cggctgtgcc cgagtccagc tctccgtcc agggctccctg cataaattta tccctgagga 600
gattctgatt catggccgag acttccggct gctcagagtt ggttttgagg ctggaggcct 660
agagcctcag gcttttcagg tgaccatggc cattgtccaa gccagagctc agagcaatca 720
agcccaacag tattcgggga taaccctgag caaggctggc cagggttctg gctccagaaa 780

```

```

accaccaatt cctctcatgg agacagcgga agactgggag actgagcgga agaagcaggc 840
agccagaggc tggagggtca gcacgggtcaa cgagagggtc gacgtagcca ccagcctccc 900
ccgttacttc tgggtcccta accgaattct ggacagttag gtcaggagag cttttggcca 960
ctttcatcag ggccgtggac cgcgcttgct ctggcatcac cctgggggca gtgatcttct 1020
ccgctgtgga ggcttctata cagccagtga ccctaacaag gaggatatca ggtgaggggag 1080
gcttgatgag aagaatcaag ggtaggtgct tcagggtaga ctccttattt cctgactccc 1140
tcccctccag agcagtggag ttgatgctcc aggcctgggca ttcagatg 1188

```

```

<210> 109
<211> 745
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7506225CB1

```

```

<400> 109
gttcctcgcc agcaggcctt ggctccttga acttttggcc gccatgtgct tcccgaaggt 60
cctctctgat gacatgaaga agctgaaggc ccgaatggag ctactcctc tacttgaaaa 120
agaaaagagat ggattacggt gccgaggcaa cagatccctt gtcccggatg ttgaggatcc 180
cgcaaccgag gagcctgggg agagcttttg tgacaagggtc atgagatggt tccaggccat 240
gctgcagcgg ctgcagacct ggtggcacgg ggttctggcc tgggtgaagg agaagggtgt 300
ggccctgggtc catgcagtgc aggcctctct gaaacagttc cagagtctct gctgctctct 360
gtcagagctc ttcattgtct ctttccagtc ctacggagcc ccacgggggg acaaggagga 420
gctgacaccc cagaagtgtc ctgaacccca atcctcaaaa tgaagatact gacaccacct 480
ttgcctccc cgtcacccgc caccacccct gaccctccct cagctgtcct gtgccccgcc 540
ctctcccgca actcaagtcc ccctgcctgg cgttcctgcc gcagctctga cctggctgct 600
gtcgccttgg catcttaata aaactgctta tacttccctg aaaaaaaaaa acaaacagg 660
gggggcgccg aactagttag cttggtaaac ccgggaatta attcccgaac cgggtacctg 720
cagggcgta caagatttta cccaa 745

```

```

<210> 110
<211> 795
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7506227CB1

```

```

<400> 110
ggaattcggc tcgaggaaag cgtggcggct gctgctagag cctttccctt taccgcaccc 60
aaggagctgg agcgacaaca acgacgtcgt ttccgtttcc accacctctt cctgttccc 120
tccttgagga cgccgtgccg ggtcagtgtt agcctccagc cctgggttggt gaaggcgaca 180
gaagtcatgg cgatgtttga gcagatgaga gccaacgtgg gcaagttgct caagggtatc 240
gacagggtacc agttcaaccc agccttcttt cagaccacgg tcaccgcca gatcctgctg 300
aaggccctca ccaacttgcc gcacacagac ttaccctgt gcaagtgcac gatcgaccag 360
gcacatcaag aagaacggcc aatccgacag attttgtacc tcggggacct gctggagacc 420
tgccatttcc aggccttctg gcaagccctg gatgaaaaca tggacctctt ggaagggtata 480
actggctttg aagactctgt ccgaaagttt atctgccatg ttgtgggtat cacttaccag 540
cacattgacc gctggctgct ggccgagatg ctccgggatc tgtcggacag ccagctaaag 600
gtgtggatga gcaaatacgg ctggagtgcc gacgagtcgg ggcagatctt catctgtagc 660
caagaagaga gcattaaacc caagaacatt gtggagaaga ttgactttga cagtgtgtcc 720
agcatcatgg cctcctccca gtaacttcag gtgtttaata aagatgtgtt gactcagcaa 780
aaaaaaaaaa aaaaa 795

```

```

<210> 111
<211> 895
<212> DNA
<213> Homo sapiens

```

```

<220>

```

<221> misc\_feature  
 <223> Incyte ID No: 3144431CB1

<400> 111  
 tacacttttct gtgccattaa cctctgcctg cagccccccg ccctccaccg agagcaccct 60  
 acagcaccag aaggaaatgt ctctgattgc aggagcgaga cttggaagcg ctggtcaaaa 120  
 ggtttaaaac agacctccac aactgcgtag cctatatcca ggaaccgcg ctgctgaagg 180  
 agaaggttcg aggtctcttt gagaagtacg tgcagcgagc agacatggcg gagatcgag 240  
 ggctgaacac agacctgcag caggagtaca cccggcagcg ggagcacctg gagaggaacc 300  
 tggccactct caagaagaag gtgggtcaagg agggcgagct gcaccgcaca gactacgtcc 360  
 gcatcatgca ggaaaatgtc tctctgatca aggaaattaa tgagctccgc agggagctga 420  
 agttcactcg gtcccaagtc tatgaccttg aagcagctct gaaactgacc aagaaagtcc 480  
 gaccacaaga agtttcagag acagaaccca gcagggacat gctcagcaca gctcccaccg 540  
 caagggttgaa tgagcaagaa gaaactggga ggatcattga aatgcagcgc ctagaaatcc 600  
 agcgctcag agaccagatc caagagcaag agcaggtcac agggttccac accctcgctg 660  
 gagttcggct tccttccctc tccaactccg aggtagactt agagggtgaag accaactgac 720  
 cccctctggg gagccatctc cagccacagc cagaagaaac acagcatgtc tgtccccaag 780  
 ccagacttgc gggttgagtc tgtatgggtc ctgcagcact gaccccagca acctctttct 840  
 cttgccctgg ggaatttggg acacagaata aagggtgttg ccacaaaaa aaaaa 895

<210> 112  
 <211> 2320  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2633315CB1

<400> 112  
 cgtccctcac accgagagtt cctgcgcgtg gggagttgga gagtttgcgt ggcgggaacg 60  
 cggcgccagc gagagcgagc ggcgcgggccc cttgcgtccg gtgcggcgat gctgaccccg 120  
 gcgttcgacc tcagccagga tccggacttc ctgactatcg ccatccgcgt gccctacgcc 180  
 cgggtctccg agttcgacgt ctacttcgag gggctctgact tcaagttcta cgccaagcca 240  
 tacttttctca gattaaccct tcttggaaag attgtagaaa atggaagtga gcaagggtcc 300  
 tatgatgcag ataaaggaat ttttaccatt cgcctgcccc aagaaacccc tggccagcat 360  
 tttgaggggg tgaacatgtt aactgctctt ctggcaccaa gaaaatccag gacagcaaaa 420  
 ccacttgttg aagaaatagg tgcttctgag attcctgagg aagtagttga cgatgaagag 480  
 tttgattggg aaattgagca gacaccctgt gaagaggtat cagaaagtgc tttgaatccg 540  
 cagtgccact atggatttgg aaacttacga tcaggagtgt tgcaacgggt acaggatgaa 600  
 ctgagtgatg ttattgatat taaggatcca gatttcaccc ctgcagctga acgaagacag 660  
 aagcgcctgg ccgtgagct ggccaagttt gatcctgatc attatctagc tgactttttt 720  
 gaagatgagg cgattgaaca gattttgaag tataatcctt ggtggactga caaatattca 780  
 aaaatgatgg ccttttttgg aaagagtcag gaacaagaaa atcatgtctac attagtgtct 840  
 ttttctgaag aagagaagta tcagctacga aaatttgcata ataaatctta tctgctggac 900  
 aagagagcct gtcgtcaagt gtgctacagt ttgattgata tccttctggc atattgctat 960  
 gaaacccggtg tcaactgaagg agagaagaat gttgaatctg catggaatat caggaaactg 1020  
 agtccaacac tatgctggtt tgagacttgg actaacgttc atgatatcat ggtgtctttt 1080  
 ggaagaagggt tgttgtgtta cccactctat cgcattttca agctgggtgat gaaggcctac 1140  
 agggacacta taaagatatt gcaactgggt aaaagtgcag ttttaaagtg tctcctggat 1200  
 attcacaanaa tttttcagga aaatgaccca gcgtacatac tgaatgatct ctacatctca 1260  
 gactactgtg tgtggattca gaaagtcaaa tccaaaaagt tggcagctct tgcagaagcc 1320  
 ttaaaggaag tctcccttac aaaggccag ctgggggttag aactggaaga actagaagca 1380  
 gcagcactgc ttgtccagga ggaagaaact gcattaaaag cagccattc agtttctggg 1440  
 cagcagacac tttgctccag ctctgaggca agtgattcgg aggactcaga cagcagcgtg 1500  
 tcatctggaa acgaagactc aggtcagat tcagaacaag atgaactcaa agatagtcca 1560  
 tctgagacag tcaattcttt gcaagggtccc tttcttgaag aaagcagtg ctttcttatt 1620  
 gttgatgggtg gtagtcgagc aaacacagcc atccaggagt ctgatgccag tcagggaag 1680  
 ccacttgctt cttcctggcc tcttgagtg tctgggctc tgatagagga gcttggggaa 1740  
 caactgaaga ctacagttca ggtttctgaa cccaaggcca ccaactgctg aaaccgcagc 1800  
 aatattcagg agagagacgg ctgtcagaca ccaaataatt gactcttagg tggttttatt 1860  
 cattgttgag aaatatggta gattgggttt catttaccga atgagaattc ttcattttca 1920  
 ctttgtaatt tttcttagta tatagtcagc ccactgtatt tgtgtgttcc acatctgtgg 1980

```

attcaaccaa ctgcagatca aaaatattga agaaaaaatc gcatctgtac caaacatgta 2040
cagacttttt tcttgttatt attctctaaa taatacagta tgacaactat ttccacagca 2100
tttacattgt attgggtaat ataagtaatc tagtgatgat ttaaactgtg caggaggatg 2160
tgggtagggt atatgcaaat actgcacat tttttatcag gggcttgagc atctgaggat 2220
tttagtatcc tcaggagtcc tggaaaccaac ccccccacaga tacgggggaca acttttatgac 2280
attgtttttc aaccaatgaa tggtataacct tttgttttcc 2320

```

&lt;210&gt; 113

&lt;211&gt; 1601

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3401751CB1

&lt;400&gt; 113

```

gtgagacagc agccggcctg gggacctggg ggagacacgg aggaccccct ggctggagct 60
gaccacacaga gtagggaatc atggctggag aattggatag cagagtaatg tttgacccct 120
ggaaacactc attctcaatg caagatcctc cgctgcaatg ctgagtacgt accgtccact 180
ctgagcctta gagggtggggg ttcacacagga gcacttcgag gaggaggagg aggaggccgg 240
gggtggagggg tgggctctgg cggcctctgt cgagccctcc gctcctatgc gctctgact 300
cggcgacccg cccgcacctg ccgcggggac ctgccttcc attcggcggg acatggcatc 360
gaagacctga tgatccagca caactgctcc cggcaggggc ctacagcccc tcccccgccc 420
cggggccccc cccttccagg cgcgggctcc ggctccctg ccccggaacc ttgtgactat 480
gaaggccggg tttcccggct gcattggctg ccccggggt tcttgcatg cgcttccttc 540
ggggaccccc atgtgcgcag ctccaccat cactttcaca catgccgtgt ccaaggagct 600
tggcctctac tggataatga ctctctcttt gtccaagcca ccagctcccc catggcgctg 660
ggggccaacg ctaccgccac ccggaagctc accatcatat ttaagaacat gcaggaatgc 720
attgatcaga aggtgtatca ggctgagggt gataatcttc ctgtagcctt tgaagatggt 780
tctatcaatg gaggtgaccg acctggggga tccagtttgt cgattcaaac tgctaaccct 840
gggaaccatg tggagatcca agctgcctac attggcaca ctataatcat tcggcagaca 900
gctgggcagc tctccttctc catcaaggta gcagaggatg tggccatggc ctctcagct 960
gaacaggacc tgcagctctg tgttgggggg tgcctccaa gtcagcgact ctctcgatca 1020
gagcgcaatc gtcgggggag tataaccatt gatactgcca gacggctgtg caaggaaggg 1080
cttccagtgg aagatgctta ctccattcc tgtgtctttg atgttttaac ttctggtgat 1140
cccaacttta ccgtggcagc tcaggcagca ctggaggatg cccgagcctt cctgccagac 1200
ttagagaagc tgcactctct cccctcagat gctgggggtc ctcttctctc agcaaccctc 1260
ttagctccac tctttcttgg gctctttgtt ctgtggcttt gcattcagta aggggacat 1320
cagtcaccat actagtttgg aaatgatttg gagatacaga ttggcataga agaattgtaac 1380
agaatcatta caggaagcag ggcctaggag cacacgtgaa acaatgacat tatccagagt 1440
cagatgaggg tgcagtcacg ggttgaaatt atcacagaat aaggattctg ggcaagggtta 1500
ctgcattccg gatctctgtg gggctcttca ccaatttttc cagcctcatt tatagtaaac 1560
caattgtcta aatccattta ctgcagaatt caccctttat a 1601

```

&lt;210&gt; 114

&lt;211&gt; 2857

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 045680CB1

&lt;400&gt; 114

```

ccacctggta ttaaacctac catccaaatc aaacaggagt cagggtgttaa aatcatcaca 60
caacagggtc aaccaagtaa aatcttacct aaaccagtga cagcaactct acccaccagt 120
agcaattccc ctattatggg ggttagcagt aatgggtgcaa ttatgacaac taaactggta 180
accactccta ctggcacaca agcaacctat acccggccaa cagtgagccc atccattggg 240
cggatggctg agccccctgg agctgcaacc tatgtgaaaa ctacgagtgg tagcatcatt 300
acagtagtac ccaaatcatt agctaccttg ggggggcaaga taattagcag taatatagtt 360
tctggaacga ctacccaaat cactacaatc ccaatgactt ccaagcccaa cgtgattggt 420
gtacaaaaga ctacaggaaa aggaacgacc attcaaggcc tcccgggcaa aaatgttgct 480

```

```

acaacgttgc taaatgctgg aggagaaaaag actattcaga cagtgccaac aggagcaaag 540
ccagctatcc ttactgctac aagacccatc accaaaatga ttgtaacgca gccaaaagga 600
ataggttcta cagttcaacc agcagctaaa atcatcccaa caaaaattgt ttatgggcag 660
caagggaaaa cgagggttct tattaaccct aaaccagtga cttttcaagc gacagttggt 720
agtgaacaaa caagacagct agtaacagaa acattacagc aagcatccag ggtagcagag 780
gctggtaatt catctattca ggaaggaaaa gaagaaccac agaattatac agatagtagt 840
tcctcttcta cagagtcctc ccagagttcc caagattccc agcctgtagt tcatgtaatt 900
gcttcccggc gtcaggattg gtcagaacat gagattgcaa tggagactag ccctaccata 960
atztatcagg atgtatccag tgaatcacia tcagctactt caacaatcaa agctctgtta 1020
gaactccaac agacaacagt aaaggaaaaa ttggaatcta aaccaagaca accactatt 1080
gacctgagtc aaatggcagt gcctattcag atgaccagg aaaagagaca ttctcctgag 1140
agtccatcaa ttgctgtggt agagtcagaa ctagtagctg aatacatcac tactgaacgc 1200
actgatgagg ggacagaggt tgcttttccc cttctagtcg gccatcgctc ccagcccaa 1260
cagccttccc agccccagcg gaccctgctc cagcatgtgg ctcagtcaca gaccgcaaca 1320
cagacttcgg tgggtggtgaa gtccatccca gcatcttccc ctggagcaat caccacatt 1380
atgcagcagg cattaaagcag tcacactgct tttaccaaac acagcgagga acttggaaact 1440
gaggaggggc aggttgaaga gatggacact ttagaccctc agacaggtct gttttaccga 1500
tctgccctga ctcagtcaca gtcagctaaa cagcagaaac ttagccagcc cccgctggaa 1560
cagactcagc tgcaagtga aactctgcag tgcttccaga ctaaacagaa gcagaccatc 1620
cacctgcagg cagaccagct ccagcacaaa ctcccgcaa tgccccagct ttccatcagg 1680
catcaaaaac tcacccctct ccagcaagaa caagcacagc ccaagccaga tgtacagcac 1740
acacagcatc ccatggtggc caaagacagg cagcttccca ccttaatggc acagcccccg 1800
caaactgtag tacagggtgct tgcagtgaac accacgcagc agctccctaa actgcagcag 1860
gctccgaacc aaccaaataa ctacgtgcaa ccccaaacc cccagagcca aatgtcgctc 1920
ccagcttctt cagagaaaca gacggcaagc caggtggagc agccaattat aaccaagga 1980
tcctctgtta caaagataac ttttgagggg cgccagcctc ccacagttac aaagataact 2040
ggtagcaggt ctgtgcctaa gctgacatca ccagttacaa gcatatctcc cattcaggcc 2100
tctgagaaga cagcagtgct tgacattttg aaaaatgtct tgatggaagc tcagattgat 2160
acaaatgtag aacatatgat agtggatccc ccaaagaagg ctcttgccac tagcatgctc 2220
actgggtgaag caggatcatt accctccacc cacatgggtg tggcagggat ggcgaattcc 2280
actccccagc aacagaaatg tagagagtcc tgttcgagtc catccaccgt tggctcttcc 2340
ctaacgacaa ggaaaattga tccaccagca gtgcctgcga caggccagtt catgcgtatt 2400
cagaatgtag gccaaaagaa agctgaagag agtcagcag aaattatcat ccaggctatt 2460
cctcagtatg ctatttcctt tcaactccagc tccaatgtgg tggtaggagc cagtgggctt 2520
cttgagctaa acaacttcac tagtcaacag ctggatgatg aggagacagc aatggagcag 2580
gacatagaca gtagcacgga ggatggaaact gaaccagcc cttctcagag ctctgctgaa 2640
cggtcctagt gtttgagcac aatagtgcac tttaaaacct gcttggttac caagtgtcca 2700
gggaaacctt tgtattttga tgactaaaaa gactcctttg cccgtactta ggctgtggac 2760
cctaaaacag cagtgtttca acaagatggt gctgcaggag cagcttttta aaacaagata 2820
aaactcacag gggaatgtac ttttttaaaa aaaaaaa 2857

```

&lt;210&gt; 115

&lt;211&gt; 1469

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1503172CB1

&lt;400&gt; 115

```

gcgagaaccc tgcaaathtt cttcctcata attgggagaa gactcactgg ccgaatggca 60
gcagtagatg acttgcaatt tgaagaattt ggcaatgcag ccacttctct gacagcaaac 120
ccagatgcca ccacagtaaa cattgaggat cctggtgaaa ccccaaaaaca tcagccagga 180
tcccaagag gctcaggaag agaagaagat gatgagttac tgggaaatga tgactctgac 240
aaaactgagt tacttgctgg acagaagaaa agctccccct tctggacatt tgaatactac 300
caaacattct ttgatgtgga cacctaccag gtctttgaca gaattaaagg atctcttttg 360
ccaatacccg ggaaaaactt tgtgaggtta tatatccgca gcaatctaga tctctatggc 420
cccttttgga tatgtgccac gttggtcttt gccatagcaa ttagtgggaa tctttccaac 480
ttcttgatcc atctgggaga gaagacgtac cattatgtgc ccgaattccg aaaagatact 540
gtggattatc cccagaaaag ctgttcgttg gattctagtc atgattgccg tgggcattctc 600
aggatctctc ttggcaatga cattttggcc agctgttcgt gaggataacc gacgcgttgc 660
attggccaca attgtgacaa ttgtgttgct ccatatgctg ctttctgtgg gctgcttggc 720

```

atactttttt	gatgcaccag	agatggacca	tctcccaaca	actacagcta	ctccaaacca	780
aacagttgct	gcagccaagt	ccagctaattg	aggaaagact	cacttgagat	accctctcct	840
tgctgaagtt	tttcttgact	tctccagttc	tcttttgttt	tttgagcat	ggttctttgg	900
gaagtggcat	ccactgcagg	aaagcagaat	gagcagagcc	agcagaactg	atggagtggc	960
acaaattccc	agtgtctgga	tgggtgccaca	ctggcgcccta	atcacccgtt	taacaagcag	1020
aaattaaatg	ttgtctagca	catgtgtctt	tcagctcttc	cttttcaccc	atggatgatc	1080
attgcgagca	tgcgctgatt	ggactgaaat	gccggggaat	aggttaggca	tgctcagtgc	1140
cgctccctttg	ccaccacagt	caaatgacat	gcttcactgt	ggtaccttaa	tacctgaaat	1200
agaacccatgg	aaaattctga	tgtcctctct	ctgaattatg	tacagactac	ctgggggatc	1260
ctcttctctc	caaatgttag	ccatcctgaa	gtagccgaac	agtagaaact	ttggtgggga	1320
ttaaccggga	gcttgaaaat	ttgtctttgg	taacctgata	ctggacagct	gaactgaatg	1380
gctgcaaat	aaatacctca	catgggaaaa	aaaaaagggg	gggggggggc	gcaattggag	1440
ccccgtttac	ggagaggatt	ttacaaagg				1469

&lt;210&gt; 116

&lt;211&gt; 4976

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1818665CB1

&lt;400&gt; 116

cccttcagcc	cagaccgagt	gttcctgtga	aagaaaatgc	cgtggataat	tgttttcagg	60
ccacaaaaat	gggagtcttc	gagtcctttt	agctatgggt	tcttcaaatac	aaataatggc	120
actacaaaga	ttaaaagaat	gaagaaggaa	cactcccttc	ccttccagcc	cttaggccag	180
cccatttcct	ggaccagcca	actgcagcat	ctgttgctat	ggcagaggac	cgaggaaatg	240
gactgatgga	gcactgcttt	gagatccata	tagagatggg	taaagggcta	gccctctctc	300
aggcaacagt	ctggggagaa	gcagattggt	atgtccagta	ctactttcca	gttcaacact	360
ctcaatccag	tgtgctgaaa	ggacctgagt	tccttgaaaa	tggaattact	ctgaagccct	420
tcagaactgc	aaccacactc	tgtgttccag	atcccatctt	taatagttaa	caccatcact	480
ctctcctgtt	gccagctgag	gttccagtgc	aaaggctcct	actaagtgct	ttctctgcac	540
agggcctcgt	gcctggaggt	ggagtccagt	ttgaaatctg	gtgcagatac	tattatccta	600
atgtgagaga	ccagaaggct	gccaaaggaa	ccttgccatt	atcaaggatc	tgtgctatgg	660
taaccaccca	gcacgtgtag	gatgtgggaa	tacagacctt	taatctccct	ttaaccccca	720
ggatttgaaa	caggaaagaa	ttgaggaacc	agtcctcagg	tttactggat	gtgggcctaa	780
ggtacaggcg	tagtccaaga	acagcagagg	gagttcttgc	tgcccgaact	gtttccatct	840
cagtccagat	tatcagagcc	tgtggtctgc	aagcagcagc	caaggctttg	gctgaacagg	900
aacccgctct	acagtttagt	gccacagtgc	gggtcaatgc	ctctgtcacc	actcatctct	960
ccttcctgcc	ccagggagaa	cagcgccgaa	cccaccctgt	ggcctgttct	ttctgccctg	1020
agttctccca	tcacgtttag	ttcacatgta	acttgggtgac	tcagcactgt	agtggagagg	1080
ctgttttctc	agcagagttg	ttggagtttg	cagaagttat	ttttgctgtc	tatcatgaaa	1140
ataccaagtc	agcaagtgat	ataatcagta	ttgagtcatg	caaagagtat	ctgcttggag	1200
tagtaaaagt	tccaacaaaa	gagctgctga	tcaagagatc	tgggatcaca	ggatgggtatc	1260
ctatcatttt	accagaagac	gggggcctac	ctcatggcct	ggagctcatg	cagaagatcg	1320
tgggtggtct	ggagctttcg	atttccttca	cgcacgtggg	agatagagaa	cgggtgtttg	1380
aagctgctga	gcatttgggc	tggagctttg	agaacagcct	gaaagatttt	gtcagaatgg	1440
atgaagggga	gccagccact	gtcaccatct	ccaccccaag	gctgtggctg	cccatccatt	1500
gtgtgctgct	tgctggccac	aaccacattc	ataagaatac	atattgctac	cttcgctaca	1560
agttctatga	tcatgaagcc	ttttggaccc	ctctcaagaa	gcctaaggaa	tctgtaaaca	1620
aaaagcagat	tatggtcact	ttcaaggcat	ccaaaagagc	agaagtcacc	aggggcccac	1680
cactgctttg	gtacttcagg	gaggagaggc	tagagatcca	agtgtggcga	gcttatggca	1740
atgacagtgt	ggagagaccc	catcagacag	acagctggat	tggctcagcc	tatgtggacc	1800
tggccagact	tggggagagg	tcagcgagga	cgtaaacctg	cagtgggtgtg	tatcctctgt	1860
ttggacgaaa	tgcttccaac	ctctcaggag	ctgccttgcg	agttcatgtg	gttctttcct	1920
ctctttcctc	acaccttgag	cccactcatg	agctggactc	catggactgc	agcagccaca	1980
gtgagtctga	gcagctcccc	agaaggaatg	atgaggtcca	gctctctcca	ccagaagtca	2040
tctcttgcca	ccagaagtct	cctgcctcca	cccaggtccc	ctgcagcagc	accacagctg	2100
aagtcgcctc	gacgcaggag	ggcctgtctg	atttggatgg	aacgtttgca	gtcagcatcc	2160
tagtagaaag	agcaatgcac	ttgagcttga	aagggagccc	cttgacagag	cggaaagtat	2220
cgatacccag	ttgtttgtga	tcctttgcaa	cagccgatga	gtcatctcct	gtatacaccc	2280
aagtggttga	aaacacagat	tccccatct	ggaattttca	acagcagtca	aggctatcaa	2340

```

aagagctgct tctggaccca caacaaaccc tgggtcttcaa agtttggcat aaaggagatg 2400
aggagagggt gattggcttt gcctcggtag acctctcccc acttctctct ggcttccagt 2460
ttgtctgtgg ccggtacaac atcacagact tcaagtggaga gtgccagggg cagataaaaag 2520
ttgctgtctc ccttttggag agtttgatag acttcaaaga agaaaggcaa gcaaggcgtg 2580
gagtggagac ctcaaaatca ctgatcccaa tatacagtcc cttttccttc cctgcctctg 2640
atacgtatgc tgcattctcc agccacatgg caaggcagac cctagaccaa cttgctcatg 2700
cctcctcaaa ggagcttgat ttctcctctc ctgggagaag tgataccaca agaagccaag 2760
catcacgcca tgaagagcat gtgcagaaca ttgcgcggtt tcatgaatcc ctgcatcttc 2820
agggagaggc acccttgcca tgtgatgaca aactgaccac atcacctttg tcctcccaaa 2880
cctccattct gacttctctc aggaagaatc tgagttagct tgatcagatt cagagggtact 2940
tccgccagaa gctcaccaag cttttcctac ccctcagccc tcagactcaa acggccatct 3000
cacagcacca ggagagctgt agggaccatc ttgggcccagg tgccagcagc ctagaccctg 3060
ggagccagtg tatcctggag aaatccagta acctggtgtt gcaagtcagc tccttaatca 3120
cagatctgca gactatcacc agggattcgc aagcagcttt gagttctcac cgagccagga 3180
gtagaagcaa caaggtccacc accctcccag atgctcagga cacagaagca ctgcaagaac 3240
gatgcacaat gccagatgag ccattggtaa gagctccaga taaaggcaca gattccccat 3300
caccctctcc tctcgaagag acttcaaata gaggaagaat gctccatgag tccctgaggc 3360
atgcagtacc cattacaagg atgcagagca gtgaagacac tgaggcaggg ccagcctaca 3420
gtgatgagga ctatgaagaa gacatcattg agcccaggac cttaaatgag atcaccacgg 3480
tgacagacaa aactagccct tgggtctagt tcatatcaga cacaagttag gtcctcagcc 3540
ctcagcctga cgagggtgcag agggaaaggcc cctcttgctc ttctccaggg cctttctgca 3600
gagaggagct tatggtcaaa agtagctttc tctccagccc agaaagggct gtcaaccccc 3660
acctgcctag gcagggttct cctagccaaa gtctgtgtgc ttgtgagtgt gaggcctcta 3720
aggccagggt tgggtggcag tctgcctcag ccaaccctca gcccatccca tgcccaacac 3780
tctcaggagc tcaacagagc agcacgtttg ttggatggag ctcaccacag acagatcaga 3840
acaaggagcc caagtcagag gcccagctg agaatgaagc tgccaccagt gagctcgggt 3900
actctgctga tagcttcaag aagttgccgc taaatctagc ctcccagagc agaaggga 3960
accataaggg accaccatc gactcctctg acatcaggca gaggcaggtg acaactggat 4020
cagagacgtc cacaagcag agcctcctgc tcccagggcc cattgtggtg cccaacttct 4080
ttttgcctcc ccagcagttg gaggttccc tgccgatgct ttcactctca gcaactttgc 4140
caccagcagc cacaacagat caggacaaga gcgaagccac caggggagct ctctcccaa 4200
ggccttgctc ccccagacct aactcgctcc ccctcaacct gcctgaggaa gaaactctca 4260
ggattgcacg gatattttct tctcagtact cccagaaaaga ctgacgcagc ctgaggggct 4320
ggggggaagg agacacctgc ccaccctccg tcagctttgt ccagcttgca ggaggacca 4380
ggtctggctc cttcagggtc gtcacagtcc tgaaaccacc acttgcctag gccacggatg 4440
cctcaagaga cccggtcagc ccagactgga ggagtgtccc aaaccagtcc agtgtccact 4500
tgccagaatt ttagctcctt ttttctgtgg tgaggcaaaag gctagaaata aaatggaatc 4560
cgcgcccaag tgggcccacaa gcaggcatct ttccggtggag aaacgcccc cagtgtcctt 4620
gacacagctg actgctggca tgagcccacc ctctgcctga agatgcagga ccatttttgc 4680
cttaaaaagt ggggaggagg gacatgtaga ttgtgtgcat tgtgcatagt caaggaaatc 4740
ctagggccac ctccagttca tttgtgtggg aacaaggata ttttatagat acaaattatt 4800
tttatgctgt gttgaattaa tcaattagga gaggaagggg aaatcacttc cttcaaactt 4860
tttatctgat tgtctaaaat tctaaccatg cttttaactt attattttta cccagctctg 4920
aaggtcattg ttcttgctg tgtttgaata aaatcatatt ggtgtttgta aaaaaa 4976

```

&lt;210&gt; 117

&lt;211&gt; 3226

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3251352CB1

&lt;400&gt; 117

```

ggccacctcc ctgttggaca ccttggctct cctacccccg cgggtagcga tctggggcga 60
ggctttgtcc ccgctcgtg aacagcgccc ctgcaggcc acggcctgcg agtgcctgcg 120
agagctagag agctgcaagc ccgggctgct ggggggctcc ctggggttgc tgccgggctc 180
gctggggcag gaaggccctg tccagccact cagcctgttg ctggccctcg ctttgcgcaa 240
caccttgggt ctccagtcct ggggtggggc ttccgtgggg ggactgctca cggataagg 300
ctccccaact ggggggtggt cctgggattg ggagactagt ggaggagggc gatggagcc 360
ttcagcccca ggcaccagc tggccggcag ctgaggaggg agagggggag cgtagcctta 420
cagcacgaga gcacagccct gaggagggcg gggagctgcg ggctgcggtg atccagcttc 480

```



tggacacctc	ctatctgctc	actcctgtgg	cccaggccca	gctcctgtgg	ctgctgggct	540
gggcccctgcg	gggtctgcag	ggacagccac	cggcactctt	caagccgcag	ctggtagggc	600
tgctaggcac	agcacagctg	acactgtttg	acgccatgct	tgcgctcaag	gcggcctttg	660
gtgaggcctt	gttcacagcc	caggatgaag	cgttgctgct	ccgccggctc	accttggtg	720
cccagcacc	tgctctgcct	ccgcccaccc	atctctttta	ccttcaactgc	gtcctgagct	780
tccttgagaa	ctggccgctg	ggccctgaag	gtgaggaggc	tgccccactg	ctgctagggc	840
cccagctatg	ccgtgggtctc	ctgcccagtc	tcctgcatga	cccaatggcc	ctcctggccc	900
gcttgcatth	actgtgcctg	ctctgtgccg	aggaggaaaga	agaggagaaa	ggccagcttc	960
caagcccacg	gcactacctg	gaagagctgc	tggctggctt	gcggcagcgg	gcagccctgg	1020
atggggggccc	ccgggccttg	gccactctct	gcttccaggc	ctcgtatctg	gtggcctgct	1080
gcctggctgg	gcaacctacg	gtgctgaccc	ccttgatcca	cggactggcc	cagctgtacc	1140
aagcccggcc	catgctggct	ccccactttg	tggacctctt	ggatcagggtg	gactctgagc	1200
tgagggagcc	cctgaagggtg	gtgttgccgg	aggtgggtgg	gtccaggccg	ggcagggatg	1260
aagctctttg	ctggcacctg	caaagtctgg	caaagggtgg	agatggagat	gcccagagtg	1320
ctaccctcaa	ctttctacag	gccgcggctg	cccactgcac	gaactgggac	ctacagcagg	1380
gcctgctgcg	ggctgcggcg	gcgctgctgc	gggcagggggt	gagggggcgg	ctggctcgact	1440
tgtgcagggt	gctggccagg	cagctggagg	accctgatgg	gcgtgaccac	gcccgcctct	1500
actacatcct	gctggcacac	ctggcagcac	ccaagtggg	ggtggccctg	ggcccctcgc	1560
ttgccgcacc	tgcactggcc	tcttcaactg	tggccgagaa	ccagggcctt	gtggcagcac	1620
tgatgggtga	ggaggcaccg	gccctggtac	ggctgagcct	ggggtcccat	cgggtcaagg	1680
gcccactccc	agtgttgaag	ctccagccgg	aggcctgga	gcccattctac	tctctggagc	1740
tgcgcttccg	tgtggaagga	cagctgtatg	cacccctgga	ggctgtccat	gtgccctgcc	1800
tgtgtcctgg	ccgcctgcc	cgccctctgc	tctgcctct	gcagccccga	tgcccgcccc	1860
ccgcacgggt	ggatgtccat	gccctttaca	ccacatccac	tggctctcacg	tgccatgccc	1920
acttgccacc	cctgttcgtg	aactttgccg	acctctttct	gcctttcccg	cagcctccag	1980
agggggccgg	gctgggcttc	tttgaggagc	tctgggattc	ctgcctgcca	gagggtgctg	2040
agagtcgtgt	gtggtgtcca	cttggggccac	agggcctgga	gggcttggtg	tcccggcacc	2100
tggagccttt	tgtggtggtg	gcccagcctc	ctaccagcta	ctgtgtagca	atccacctgc	2160
ccccggactc	aaagctgctg	ctgcggctgg	aggcggccct	ggcagatgga	gtgcctgtgg	2220
ccctgcggac	cgatgactgg	gccgtgctgc	ccctggcggg	ggactacctc	cgtgggctgg	2280
cggtgctgt	ctgagccccg	ggagaccagg	tgggggcagg	actgtggccc	ttgtgggggc	2340
caaggcacac	tcctgtagct	ctgtcgccaa	aacctgcat	tccgcagtgc	cctcgctggc	2400
ttgttttctt	ttgggccccg	gttgggagca	ggctcctggg	ggtgagggtc	tgtctgagtc	2460
tgttttttgct	gctctagcaa	gatccctgag	acggggtaag	ttataataaaa	cagaaatgta	2520
ttggctcaga	gctctggagg	ctgggaagtc	catagtggag	ggcccagagc	ccagtgagga	2580
cccttctgca	tcttcccatg	acagaggaca	gagtgcacag	agtgagggga	gagagcgaac	2640
ccactcccaa	gataacggcc	tgagtccacg	caggagggca	gagccctcgt	ggcctgacca	2700
ctcttaggg	tctaattcta	tactgttaca	gtggcaatcg	catttcaacg	tgagcgtggg	2760
aggggacaaa	cattcaaac	atagcagggt	cataggcccc	aagaaggctt	ccagacagaa	2820
cccggcactg	tagagctaca	gagttccttc	tcaccagac	acactagtgc	tctccagaat	2880
ctctgcccag	gaggcagggg	aaaaatctcc	tttctgcttc	ccagagaggg	tctggccttg	2940
cccaaggcca	ctcagtgatt	taggaagaag	gctgaagcta	gagtgtggac	ctgacctgtt	3000
ggacagtact	tttctttctt	tttttttcgg	ggaggggtta	aagacgggat	ctcgtctctg	3060
cgctctggag	tgcagtgggtg	tgattagggg	tcattgcaga	ctctgcctct	caggctcggg	3120
tgatcctccc	atctcagcct	gcctccgagt	agctgggact	acagggtgtga	gtcaccactg	3180
tatttttgca	gacaggatct	ccctgtttct	aggctggtct	cgaatc		3226

&lt;210&gt; 118

&lt;211&gt; 1260

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 55091643CB1

&lt;400&gt; 118

gccgggccctg	gctgagaggc	gttaggagtc	cgggggttcg	ccgcgcggag	ccggggagca	60
gccgaccatg	gagccccaga	aaatcatgcc	accctcaaa	cctcatccac	ctgtcgtggg	120
caaagtgact	catcacagca	ttgaattata	ctgggatctg	gaaaagaaa	ccaaacgcca	180
aggacctcaa	gagcagtgg	tcagcttctc	gattgaagaa	gaagacccca	aaatgcacac	240
ttatgggtatc	atattatacgg	gatatgcaac	gaagcatgtt	gttgaagggtc	tggaaaccaag	300
gacgctgtac	agatttcgcc	tgaaggctac	cagccccctc	ggggagtgtg	agtacagccc	360

```

actcgtctca gtgtctacaa ccagagagcc cataagtagt gagcacttgc accgggctgt 420
cagtgtgaat gatgaagatt tgctgggtccg aatacttcaa ggaggccctg ttaagggtga 480
tgttcccaat aagtttggct ttaccgctct gatggttgc gccagaaag gatacaccag 540
gcttgtgaaa atcctagttt ctaatggcac agacgtgaat ctgaagaatg gaagtggcaa 600
ggacagtcta atgctggcgt gctatgcggg acacctagat gttgtgaaat atctccgaag 660
acatggcgct tcttggcagg ctagagacct gggaggctgt acagctctgc actgggctgc 720
agatggaggc cactgcagtg tgattgagtg gatgataaag gatggctgtg aggtagacgt 780
cgtggacact ggttcaggat ggacccccact catgagagtc tctgcggtgt cgggaaatca 840
gagggtggcc tctcttctaa ttgatgctgg ggccaatgtg aatgtgaagg acagaaatgg 900
aaagacgccc cttatgggtg ctgtgtttaa taatcatgaa gagttagttc agttacttct 960
tgacaaaggg gcagatgcaa gtgtaaaaaa tgagttcggc aaagggtgtc tagaaatggc 1020
cagagttttt gacagacaga gtgtagtctc cttattagaa gaaaggaaaa aaaagcagag 1080
gccaaagaag tcttgtgtct gctgatgaga gcaccactca tctgcgaaac gcacgtaaaa 1140
caaagtgaac cgtgactgtt aaactagggg tgggaaattc tgcactcttg ggtgctgtac 1200
atattttat ttagttgaag attcactgat ccactttga aatacatctt tttacctaa 1260

```

<210> 119

<211> 649

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7500770CB1

<400> 119

```

ggcggagctg agggaccccg cggctggtga gcgcccgtg gaggctggag cttccggggc 60
ctggaaaggg ggtcccccg cgccccgggt cggaggcaga cccctgggtt tgggggacat 120
gggcatttgg ggcgcctgaa cccaagacct ctggatgagc tgccccgttc agaccatgga 180
tcctgaggtg accttgctgc tgcagcccc agcaccagga cctcgctggg cagctgggtg 240
tacatgaact cttttccagt gtcttcagg agatctgtga tgagggtgaac ctgccgctgc 300
tcacctgag ccagccccg ctggtgggca tcgcccga aa tgagaccagt gctggccgag 360
ccagtgcga gttctatgtc cagtgcagc tgacttctga gcagggtgagg aagcactacc 420
tgagtggggg acccgaggcc cacgagtcta caggaatctt ctttgtggag acacagaacg 480
tgcgagatt gcccgagacg gagatgtggg ctgaactctg cccctcggcc aaaggcgcca 540
tcactctcta caaccgggtt cagggaagtc ccactggagc ggccctaggg tccccagccc 600
tactccccgc gctctgaaaa taataaacga ctttattctt ggaaaaaaa 649

```

<210> 120

<211> 3439

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7506350CB1

<400> 120

```

ggttggcgct cggctcctgcc ctgcgcgttt tgggctgcgg ttggaggggc gcagttgtag 60
tcagggtccg ccagctcggg gccagcgcat ggggctgctg agaccgctcc ggacgtgcga 120
agcgttcgag gtgcggtagg aagctgaagt cttcagcaga accctgacct aggtaggagg 180
tctcaatggc cctgggtgaa gaaaaggcag aagcggaagc atctgaagac acaaaggccc 240
agtcttatgg gagaggggagc tgcaggggagc gggagctgga catcccaggg cccatgagtg 300
gggagcagcc cccacgcctg gaagctgagg gagggctcat ctcccctgta tggggggcag 360
aagggatacc tgcccctact tgctggattg ggactgacct tggcgggccc tctagagccc 420
accagccaca ggccagtgat gccaacagag agcccgtagc tgagagggtc gagcctgcac 480
tcagtggcct gcctcctgcc accatggggg ctggagacct tctgctctcc ggggaaagcc 540
aggtggagaa gaccaagctt tcttctctcc aggagttccc tcagactctg agccttccca 600
gaacaacaac tatttgetca ggacatgatg ctgataccga agatgatcca tccctagcag 660
atattgcccc ggcactggac ctacagccag agcctcacag ctcaggctct tcttgctgt 720
cactgtggaa gtccgtgctg agcccagggt ccgcagctca gccttccagc tgcagcatct 780
ctgcttcctc cacaggcagc agtctccagg gtcaccagga gagggcggag cctcgtggtg 840
gttctctggc caaggctctc tctcctctgg agccggctcg ccccaggaa ccttctctg 900

```

```

tgggtggggct aggacctcgg cccagtggt caccacagcc tgtgttctct gggggtgatg 960
cttctgggct aggcaggaga cgctctctct tccagctga gtactgggccc tgtgtgctgc 1020
cagattccct gctccatca cccgaccgc actccctct ctggaacca aataaagagt 1080
atgaagatct gcttgactat acttaccac tgaggcccg gctcagctc ccaaagcacc 1140
ttgatagccg tgtgccagct gacctgtcc tgcaggactc cggggtagac ctggatagct 1200
tctctgtctc tccagcaagc accctcaaat cacctactaa tgtctcccc aactgcccac 1260
cagcagaggc cactgccctg ccatTTTTctg ggcccagaga gccaagcctt aagcagtggc 1320
cctccagagt accccagaaa cagggtggca tgggcttggc atcttggagc caacttgcat 1380
ctacccccag agccccaggc agtagggatg ctcgttggga gcgcagagag ccagccctga 1440
ggggtgcgaa ggaccggctg actataggca agcaccttga tatgggctct cccagctaa 1500
ggacacggga cagaggggtg ccctcgcca ggccagagag ggagaagagg accagccaga 1560
gtgcccggcg cctacctgc acagagtcta ggtggaaatc agaagaggaa gtggaaagt 1620
atgacgagta tcttgccctc cccgctcggc tgacacaggt ttctagcctg gtttcgtatc 1680
taggatccat ttctacctg gttaccctgc ccatgggga tatcaaaggg cagagcccct 1740
tggaagtgtc agacagtga gggccagctt ccttcccttc aagctccagc caaagccctga 1800
ttcccaggaa agggaggaga cagggaagg aatcactggt gcaatgtgtg aagacatttt 1860
gctgtcagct ggaagagctg atctgctggc tgtataatgt tgcagatgtt actgaccacg 1920
ggactgcagc caggccaat ctacaagtc tcaagtcttc tctgcagctt taccggcaat 1980
ttaagaaaga tatagatgaa catcagtctc tgacggagag tgtcttacag aagggggaga 2040
ttcttcttca gtgctgttg gagaacccc cagttttaga ggatgtcctt gggaggatcg 2100
caaagcagtc tggtagctg gagagccacg cagatcgct gtatgactct atcttggcct 2160
ctctggacat gctggctggc tgcacctta tccctgacaa aaagcccatg gcggcaatgg 2220
agcaccatg tgaaggggt taacctggca tcttctcagg cagaggtgca acctgtcggc 2280
caaacactgg ctagttagaa aaaaatcctt caagaagtc ttgcaagagc cattcaacac 2340
tgaagtcaag ggggaaacct tcaatatact agctttaaata tgcttttctc ctccctgaat 2400
ggagggtgca gttactattt ttatttacia tgaatccttc caaactaagg ttttctttt 2460
gggtgcttct cacaagagc gtgtagcttt gcaatttttg ataaaataag acgtgttgcc 2520
ttccctctca tccattaaa aaagagcaga taaagctcaa actattccac ctggtagggt 2580
catttaccaa gggctggtag gagattccga atactcaata ttctcaggca tgcaagagga 2640
cctaagagac tgaaaaaagt gaattgttg ccagaaacct tgggtctttt cataaaagga 2700
aatttgggct tgtcaatccc tctttaaaga tataggtaga aagaagagat ttttaaccct 2760
ttgtctgcaa aagaaaaagg gttaaatgaa gcatctatac agatgacttt gggagattat 2820
gacaccttt gccatcttca gggcttacgt ctttactttc ttggctaacc agtttcttag 2880
aagaaaatgt gtcagggact tgggatcta cagtcatgct ttagcatggc tatggagcta 2940
attatcaagc ttaaagggt actttgggag gactcctccc ttcactccta gtctcccttg 3000
gaagagcagt ccaggctcag ggaagggaat ggataggatg acacagtaag tacagaaaat 3060
gaagctgtca atagtgaagg aaaaaggga attcttctgt gctttggcat tgacacataa 3120
ttgtctgcat taaaaaaa aaaaactat ggatcaacca ttcaaagcat caattctata 3180
gctttttgca gcaagatgat cagttattgc tggttaactgc agattctaca gaaaagcact 3240
tttaaagttc tgttgggcca tgaatgaacc tgggaaggag aaagcacagt aacaatggaa 3300
actaaagtc tattattaat gtgacttttc tgcttcagaa tttctgagac tttatagtca 3360
ctggaaaata gccattaaat tctttaaaga ttagctcttc aacagtagct acacataaaa 3420
tgcagtattt caagtgaat 3439

```

<210> 121

<211> 1570

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7508370CB1

<400> 121

```

ggcaactggg ctggaccgaa accggcgcg agcaactgag gcccagacct tctcgggacc 60
cgggggacgc ctaaccccg gagaaccctg caaattttct tctcataat tgggagaaga 120
ctcactggcc gaatggcagc agtagatgac ttgcaatttg aagaatttg caatgcagcc 180
acttctctga cagcaaacc agatgccacc acagtaaca ttgaggatcc tggtgaaacc 240
ccaaaacatc agccaggatc cccaagaggc tcaggaagag aagaagatga tgagttactg 300
ggaaatgatg aacttgacaa aactgagtta ctgtggac agaagaaaag ctccccttc 360
tggaacattg aatactacca aacattcttt gatgtggaca cctaccaggc cccttttggg 420
tatgtgccac gttggctctt gccatagcaa ttagtgggaa tctttccaac ttcttgatcc 480
atctgggaga gaagacgtac cattatgtgc ccgaattccg aaaagtgtcc atagcagcta 540

```

```

ccatcatcta tgcctatgcc tggctgggtc ctcttgcact ctgggggtttc ctcatgtgga 600
gaaacagcaa agttatgaac atcgtctcct attcatttct ggagattgtg tgtgtctatg 660
gatattccct cttcatttat atccccaccg caatactgtg gattatcccc cagaaagctg 720
ttcgttggat tctagtcacg attgccctgg gcatctcagg atctctcttg gcaatgacat 780
tttggccagc tgttcgtgag gataaccgac gcgttgcatt ggccacaatt gtgacaattg 840
tgttgctcca tatgctgctt tctgtgggct gcttggcata cttttttgat gcaccagaga 900
tggaccatct cccaacaact acagctactc caaaccaaac agttgctgca gccaaagtcca 960
gctaattgagg aaagactcac ttgagatacc ctctccttgc tgaagttttt cttgacttct 1020
ccagttctct tttgtttttt ggagcatggg tctttgggaa gtggcatcca ctgcaggaaa 1080
gcagaatgag cagagccagc agaactgatg gagtggcaca aattcccagt gtctggatgg 1140
tgccacactg gcgcctaata acccgtttaa caagcagaaa ttaaatgttg ctgagcacat 1200
gtgtctttca gctcttcctt ttcacccatg gatgatcatt gcgagcatgc gctgattgga 1260
ctgaaatgcc ggggaatagg ttaggcatgc tcagtgccgt cccttttgcca ccacagtcaa 1320
atgacatgct tctactgtgg accttaatac ctgaaataga accatggaaa attctgatgt 1380
cctctctctg aattatgtac agactacctg ggggatactc ttctctccaa atgttagcca 1440
tctgaagta gccgaacagt agaaactttg gtggggatta accgggagct tgaaaatttg 1500
tctttggtaa cctgatactg gacaactgaa ctgaatggct gccaaataaa tacctcacat 1560
gaaaaaaaaa 1570

```

<210> 122

<211> 2671

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 2894093CB1

<400> 122

```

tagaagaacg tagcccgacc tgcgaggcca gctcgtcct gccctcctcc agttgtgcag 60
ctttcctgct gaaaagtcct aaaatttgac cgggggactg ggaaccccg cccaagctg 120
ccgcgcgctc tgaaactccg ttgctgtgga gacggaggcg gttgaccaa ggattccgaa 180
cgggtgttggg ggaggggagc cgacgtggtt caggaactgg gcgtggcttc ggttactgtg 240
gtagcgggga cgcttaggca gggcctgcgc ccagtttaga ggccaatgct tctcccgctc 300
tttgcatggc acccaagaag gagaaaggag gaactgtgaa caccagttct aagatatggg 360
aaccctcgct catagctgca cagttcaatc agaacgattg gcaggcctcc attgcttttg 420
tggttgggaa ccagattgaa gatgatcttc tcatccaagc ccttaccgtg gctgtccagg 480
tccctcagcg taaactcttc agtatggtgt cgtggcaaga cattctccag cagattaatg 540
aaataaatac acttgttggg tctgcttcat ctaaaaaggc aaaaaaacct gtaggtggta 600
atgctccttt atattatgag gctccacaac ctgctgtacc agctcctgga aagaagaaag 660
cacagtatga agaaccgcaa gctccaccac cagtgaactt agtcatcaca actgaagtag 720
acatagata ttacaattat ttgctgaatc caattcgaga ggaattcatt tctgtgcccc 780
tgatactgca ttgtatgctg gaacaggttg ttgcaactga agaagatctc gtcccaccta 840
gtctgcggga gccatcccc cagagcagac ggctagacca cagaatcgca gctcacattg 900
tgtcccttct gccctcactc tgtctctcag agagggagaa aaagaatctt catgacatat 960
ttttatctga agaagaaaat gaaagcaaa cagtgcccaa agggccccct cttactgaact 1020
atcatgatgc acacgcccac aagaagtacg cactacagga ccaaaaagaat tttgatccag 1080
ttcaaattga gcaggagatg cagtcacagt tgccactgtg ggaatttctt caattccctc 1140
tccccccacc atgggaacaac actaaacgtc tagctacaat tcatgagctt atgcactttt 1200
gtacgagtga tgtcttgagc tggaaatgaag tagaacgagc cttcaagggtg tttacttttg 1260
agagcctgaa gctctctgag gttgatgaaa aagggaaact gaaaccttct gggatgatgt 1320
gtgggtcaga ttctgaaatg ttcaacatac catgggacaa ccctgccaga tttgctaaac 1380
agataaggga gcaatatgtc atgaaaatga atactcaaga ggccaagcag aaagcagata 1440
ttaaaatcaa agacagaaca ctatttgtgg atcagaattt gtcaatgtct gtgcaagata 1500
atgaaagcaa ccgagaacct tcagatccta gtcagtgtga tgctaacaat atgaagcatt 1560
ctgacttgaa taatctcaaa ctctcagtc ctgataatag acagctgtta gagcaggaga 1620
gcatcatgaa ggctcagccc caacatgagt ctctggagca gaccacaaac aatgagatca 1680
aagatgatgc agtcacaaaag gctgattctc atgaaaagaa acccaagaag atgatggtgg 1740
acgcagattt agaggacata aagaaaacac agcagcgagc tctaattggac tggagtttta 1800
ctgaacattt taataccgaaa gtactgcttc aggtccttca agaagcccat aagcaatata 1860
gggtgtgttg ttcttactac cacacccaag acaactcttt acttttagtc tttcacaatc 1920
caatgaatag acaacgtttg cattgtgaat attggaacat tgctctccac tccaatgttg 1980
gattcaggaa ttatttggaa cttgttgcaa aatctattca agattggatt acaaaaagaag 2040

```

```

aagctatata tcaggaatct aaaatgaatg agaaaaatcat caggaccaga gctgagctgg 2100
aattgaaatc ttctgcta atcagcaaaa cttctgctag caaaattttt tccattaaag 2160
aatctaaaag taacaaagga atcagcaaaa cagagatatc agatcaagaa aaagaaaaag 2220
agaaggaaaa gattcctttc attttagaag gctctctcaa ggcatggaaa gaagagcaac 2280
atcgattagc agaagaggag cgcttaaggg aagaaaaagaa agcagagaag aagggtaaag 2340
aagctggtaa aaagaaaggc aaggataacg cagagaaaga ggatagtagg tctttgaaga 2400
aaaaatcacc ttacaaggag aaatctaaag aagaacaagt caagatccaa gaagtaacag 2460
aagagtcccc ccaccaacca gaacctaaaga taacttacct ggcccaactt ttatcaaagt 2520
gagagtggta aaggacaacc acaattttat gattcattta aatgacccta aggaaattgt 2580
gaaaaaggaa gagaaagggg attattattt agaagaggaa gaagaaggag atgaggaaca 2640
aagtcttgaa acggaagtat cagatgcaaa g 2671

```

<210> 123

<211> 3376

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7507335CB1

<220>

<221> unsure

<222> (1) ... (3376)

<223> a, t, c, g, or other

<400> 123

```

agcctggact caagaggata gactgggcaa aagcgcttct gggaatcgta gtttagggcg 60
ccgacctcaa gatccaggct gggcacgcta cccaaccaac tccttacttg agctcttcag 120
ccgaatacat cccgaaggaa atgccctgca gccgccgcca ggcatgttc ttggagatca 180
acatcctcca ggtccgtttt gaattccgac accccaagag acgttccact caccacctga 240
ctattcttaa ttcaacctca agcccggagt caccacgcca ttcaacgtgc gcttgcgcgc 300
ggaagcggtc gcaggaacga catttacgtg gatgagatca cttctacatg gggggggaac 360
aaatagttgt acgactctct gctcgctctt ttgttttgaa aacaatcaaa gatataaag 420
acatcaaaaat aatcatttag gaaaattcct ccgaggcaaa tcgggtgtat agagctttta 480
gcttccgtga cgaagtcttt caaacatga gcttgcaacc tcccatttta tgattgcttt 540
cgtcttcttc cgcttcagcc aatagcggag ctgttccctg ctccccgncc actgcgcccc 600
acttcttgaa gttcgaggca cgctttccca acatgctctg cagagaaatc aaagatggcg 660
gttgatctct ctgttcgctg gctgggcctc cgcagcaggc ttggccagcc gctgacgggt 720
cggcgggagg gtttgtgtga acaggcacgc agctgcagat ttatttctgg tagtgcaacc 780
ctctcaaaagg ttgaaggaaac tgatgtaaca gggattgaag aagtagtaat tccaaaaaag 840
aaaacttggg ataaagtagc cgttcttcag gcacttgcac ccacagtaaa cagggatacc 900
acagctgtgc cttatgtgtt tcaagatgat ccttacctta tgccagcatc atctttggaa 960
tctcgttcat ttttactggc aaagaaatcc ggggagaatg tggccaagt tattattaat 1020
tcatacccca aatattttca gaaggacata gctgaacctc atataccgtg tttaatgcct 1080
gagtactttg aacctcagat caaagacata agtgaagccg ccctgaagga acgaattgag 1140
ctcagaaaaag tcaaagcctc tgtggacatg tttgatcagc ttttgcaagc aggaaccact 1200
gtgtctcttg aaacaacaaa tagtctcttg gatttattgt gttactatgg tgaccaggag 1260
ccctcaactg attaccattt tcaacaaact gagcagtcag aagcatttga agaggaaaa 1320
gatgagacat ctaggaggaa agctgggtcat cagtttggag ttacatggcg agcaaaaaac 1380
aacgctgaga gaatcttttc tctaagtcca gagaaaaatg aacattccta ttgcacaatg 1440
atccgaggaa tgggtgaagca ccgagcttat gagcaggcat taaacttgta cactgagtta 1500
ctaaacaaca gactccatgc tgatgtatag acatttaatg cattgattga agcaacagta 1560
tgtgcgataa atgagaaatt tgaggaaaaa tactggagct gctaagacac 1620
atggttgcac agaagggtgaa accaaatctt cagactttta ataccattct gaaatgtctc 1680
cgaagatttc atgtgtttgc aagatcgcca gccttacagg ttttacgtga aatgaaagcc 1740
attggaatag aaccctcgct tgcaacatat caccatatta ttgcctgtt tgatcaacct 1800
ggagaccctt taaagagatc atccttcac atttatgata taatgaatga attaatggga 1860
aagagatttt ctccaaagga cccgatgat gataagtttt ttcagtcagc catgagcata 1920
tgctcatctc tcagagatct agaacttgcc taccagtagc atggcctttt aaaaaccgga 1980
gacaactgga aattcattgg acctgatcaa catcgtaatt tctattattc caagttcttc 2040
gatttgattt gtctaattgga acaaattgat gttaccttga agtggtatga ggacctgata 2100
ccttcagcct actttcccca ctcccaaaaca atgatacatc ttctccaagc attggatgtg 2160

```

```

gccaatcggc tagaagtgat tcctaaaatt tggaaagata gtaaagaata tggtcatact 2220
ttccgcagtg acctgagaga agagatcctg atgctcatgg caaggggacaa gcacccacca 2280
gagcttcagg tggcatttgc tgactgtgct gctgatatca aatctgcgta tgaaagccaa 2340
cccatcagac agactgctca ggattggcca gccacctctc tcaactgtat agctatcctc 2400
tttttaaggg ctggggagaac tcaggaagcc tggaaaatgt tggggctttt caggaagcat 2460
aataagattc ctagaagtga gttgctgaat gagcttatgg acagtgcaaa agtgtctaac 2520
agcccttccc aggccattga agtagtagag ctggcaagtg ccttcagctt acctatttgt 2580
gagggcctca cccagagagt aatgagtgat tttgcaatca accaggaaca aaaggaagcc 2640
ctaagtaatc taactgcatt gaccagtgc agtgatactg acagcagcag tgacagcgac 2700
agtgcacca gtgaaggcaa atgaaagtgg agattcagga gcagcaatgg tctcaccata 2760
gctgctggaa tcacacctga gaactgagat ataccaatat ttaacattgt tacaagaag 2820
aaaagataca gatttgggtga atttggtact gtgaggtaca gtcagtacac agctgactta 2880
tgtagattta agctgctaata atgctactta accatctatt aatgcaccat taaaggctta 2940
gcatttaagt agcaacattg cggtttttcag acacatgggtg aggtccatgg ctcttgtcat 3000
caggataagc ctgcacacct agagtgtcgg tgagctgacc tcacgatgct gtccctcgtgc 3060
gattgccctc tcctgctgct ggacttctgc ctttggtggc ctgatgtgct gctgtgtgct 3120
tggtccttca tcttaggtgt tcatgcagtt ctaacacagt tgggggttggg tcaatagttt 3180
cccaatttca ggataatttcg atgtcagaaa taacgcacat taggaatgac taaacaagat 3240
aatggcagtt taggctgcac aactggtaaa atgactgtag ataaatgttg taattagtgt 3300
acacgtttgt atttttgtta atatagccgc tgccatagtt ttctaacttg aacagcaaaa 3360
aaaaaaaaa aaaaaa 3376

```

<210> 124

<211> 2198

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7509081CB1

<400> 124

```

aacggccacc gcagctgcct tcggcgccgc cagccccgcc atggctgcgc cctgcgcgga 60
ggacccctcg ctggaaaggg attttaaggg ccaccgagat gcagttacct gtgtggactt 120
cagtatcaac acaaagcagc tggccagtg ctccatggac tcatgcctca tggctctggca 180
catgaagccg cagtcacgcg cctaccgctt cactggccac aaggatgccg tcacctgtgt 240
gaacttctct ccttcgggac acctgcttgc ttccggctcc cgagacaaga ctgtccgcag 300
ctgggttacc aatgtcaaaag gtgagtcac tggttttcgt gcacacacag ccacagttag 360
gagtgtccac ttctgcagtg atggccagtc cttcgtgaca gcctctgacg acaagacagt 420
caaagtgtgg gcaactcatc gccagaaatt cctgttctcc ctgagccagc atatcaactg 480
ggtccgctgt gccaaagttct cccccgacgg gcggtcctatc gtgtctgcca gtgatgacaa 540
gactgttaag ctgtgggaca agagcagccg ggaatgtgtc cactcgtatt gtgagcatgg 600
cggctttgtc acctatgtgg acttccaccc cagtgggacg tgcattgccg ttgccggcat 660
ggacaacaca gtgaagggtg gggacgtgcg gactcacccg ctgctgcagc attatcagtt 720
gcacagtgca gcagtgaacg ggctctcttt ccacccgtcg ggaaactacc tgatcacagc 780
ctccagtgc tcaaccctga agatcctgga cctgatggag ggccggctgc tctacacact 840
ccacgggcat cagggaccag ccaccactgt tgcttttca agaacggggg agtattttgc 900
ttctggaggc tctgatgaac aagtgatggt ttggaagagt aactttgata ttgttgatca 960
tggagaagtc acgaaagtgc cgaggccccc agccacactg gccagctcca tggggaatct 1020
gacagtctcc attctggagc agcggttgac actgacagaa gacaagctga agcagtgtct 1080
ggagaaccag cagctaata tgcagagagc aacaccatga tcaggggagc aggaatcagg 1140
agctcgggtg atttgacagt ggcaggccag ggatttgtac catgggactt gggtaaataa 1200
aggggactga actctgtggg aatcacatcc atactggagc cctggatttt tgcagttctg 1260
ccctccacct tgctatctgc accaggaggg cctccacctg gcagccagag gtccccagt 1320
ggccgggctc acacacaaat gatgcttcag acccgaatga gaggaccaca ttttgcttaa 1380
tgtaaaaggag ccacttgaaa atgtctgctc cttcgggggtc ctgagattgt ggctccccct 1440
ctggaggagg tggctccacg atgccttgat tttcactcat catttggaac tgtgactggc 1500
ttttcctacc tctgccatgg tgtagaaatt gattgcacat tgattggatg agccgggggt 1560
tttctctaaa tctgactaaa ggcccaaagt gggcccatct gagttaggtt tgttgagaac 1620
aagccctctc aagtgggtgg tggcttttca gtggccctga tttctgttcc acacgtgttc 1680
actggagcca ggtgacttcc tccttgctg agtgagggcc acaggaatct caaaattaaa 1740
acctgacttc attggcaaaa aaaaaaaaaa aaaaaagggg gggcgcacga acccccgag 1800
gaatcggagc acattatccc agggcctatg gccccggggg ccctcaaaaa aaaacacaca 1860

```

```

atcgaactcg gggccggaaa ctacgcgaat attcgctaga ccccccatat gtggaaagtc 1920
attggcgcg gcaagtgttt ggaaaagcgg cgttgaaggg gccaggaaat caaaacaaga 1980
cattgaacac gagccacggg ggggcccgcg acaaaagcgc acaagagaga ataagaacac 2040
cacggggcgc gcaaaacact gcgggaaacg gggccaaatc gagacaacaa acaaaaagaa 2100
gcaaaaacac gaaaggaaac aacaaataga aagaacaaca agcacatcaa aagcaaacta 2160
caaaataaac aaaaacaatg aacgaagcac gaacagaa 2198

```

<210> 125

<211> 1453

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7502450CB1

<400> 125

```

ctcctcccc tccactgac aactgcccc actgctcttc ccgccccggt cacagtgaag 60
atgtagacgg ggtcgttgtc cgtacgactg tgcgccaggg ctcggggagg ggcgccctcc 120
gcgtgagcgc cccctgga atattgaaca taatcacctc tcattccaga ctatgttagg 180
tcttaatggt gggaggacgc ccgagtgtc ggcccgtttc accccgagga ggaaggacac 240
tggtcatga cgccatcaga gggcgccaga gcagggaccg gacgcgagtt ggagatgttg 300
gactcgctgt tggccttggg cggcctggtg atgcatgggg attccgtgga gtgggagggg 360
cgcagtctct tgaaggcgct tgtcaagaaa tctgcaactgt gtggggagca agtgcatatc 420
ctgggctgtg aagtgaagcga ggaagagttt cgtgaaggtt ttgactctga tatcaacaat 480
cggctggttt accatgactt cttcagagac cctctcaact ggtcaaaaac tgaggaggcc 540
tttctgggg ggccgctggg agccttgaga gccatgtgca agaggacaga tcctgttcct 600
gtcaccattg ctctcgattc actcagctgg ctgctacttc gccttccctg caccacactc 660
tgccaggtcc tgcattgtgt gagccatcag gactcttgtc ctggtgacag ctctcagtg 720
gggaaagtga gtgtgctggg ctgtctacat gaagagcttc atggaccagg ccctgtggga 780
gctctcagca gccttgctca gactgaggtg accctgggag gtaccatggg ccaggcctcg 840
gcccacatcc tgtgtcggag gcccgcagac cgcccaactg accagactca gtggttctcc 900
atccttcggg acttcagcct ggatctccaa gaggggccc ctgtagagtc ccagccctac 960
tccgatcctc atatacccc ggtggatccc acaactcatt tgacctttaa ccttcacctg 1020
tccaagaaag agagagaagc cagagatagc ctgatcctgc ccttccagtt cagttctgaa 1080
aaacagcagg ctctcctgcg gctaggcca gggcaggcta ccagccacat cttctatgag 1140
ccagatgctt atgatgacct ggaccaagaa gaccagatg acgacctgga tatttgactg 1200
gccagatttg attgattgt aattggagg ggcgcgga gactttgggc ccagaacct 1260
ctttctattg tttgtgttag ccttaccctg tccctgcccc accttggttc cccttgtcta 1320
tgagagcccc ccttgtagc caggaagcag cgtctcatca ggacagaagg taggatgaag 1380
acatggggta atgtgagaga gtagaacacc cccgtaccta ataaaaatct ttattttttt 1440
attaaaaaaa aaa 1453

```

<210> 126

<211> 949

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7501405CB1

<220>

<221> unsure

<222> (1) ... (949)

<223> a, t, c, g, or other

<400> 126

```

cttctctct cccaccggc cttcaccttt tgttcctat cctgggcccag ttctctcgca 60
gggtccagat gtccagttcc agatgcctgg acccagagtg tgggggaaat atctctggag 120
aagccctcac tccaaaggct gtccaggcgc aatatctcat tcgtcggttc ggatccaagc 180
cagttctgtg gtcagcaagg ctccccctctg ggcaggcccc ctggtcagag ggagtttgta 240
tcctcaggga ggagtttgcg gctgaccttc cgcacacagc cttctctgga gaacaagact 300

```

```

gccacacctcc acaaggggctt cctgggccctc taccaaaccg tgggtgagtg tccctcctgg 360
gggtgcagggg aggggagcctc tgttcccagc catgaccctg gtatcttcaa gccttaagtg 420
gaagcttgag tgacagctga ggctggggac tcagggacac ctgggctgga tcccagccct 480
gcccctgctg gcaagcaacc ctattaagag acagccgtag ctgagcccc agcgggtgtt 540
tccatgcaga tttacaggcc cagtgtttgc agatcatctc atttttaaag agatgcccc 600
aatccagatt ttttaagtaa attataaatt ttcaaagat gactcagatt tgtttttgaa 660
agataaatac tggctggagg cagtgggttca tgcctgtaat cccagcactt tgggaaggcca 720
aggctggaag attgcttgag gccagtagtt ggagaccagc ctgggcaaca tagtgagacc 780
ccgtntccac aaaaaaaatt aaaaattagc aggcattggt gtgcccactg tagtaccaga 840
tactcaggac gagggggggg tgcctgaccg ggggtgggct ggggggttaac cgggtccagga 900
ataacccggt aagggtccgc cgcccatgtg gatcggtgcc ccgtctgtt 949

```

&lt;210&gt; 127

&lt;211&gt; 1265

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7504528CB1

&lt;400&gt; 127

```

gcagagtatc cagaggggccc gcatgtggtc tgcagaagag gaggacgtgg ctgggtggca 60
gggctgggtct ccaaggacga gtaagatcct gcagctgcgg gcttcaggaa gtctcctggg 120
gctatcagat ggctccttct tgtagcagca gctgtggggg cactggccc tgagccctca 180
gagggggcggc cgtgggggac ctctgtctt ttgccttgca agggcctcag ttgtgctttt 240
tccctctagg cagccatggg tgccaggcag tgcagagagc agtggggcat ggctgcagcc 300
ctgcaggtcc tgccccgctt ggcccagacc cccttgcac cactcctctg gcggggctca 360
gtggcccgtc tggccagcag catggccttg gcagagcagg ccaggcagct gtttgagagt 420
gctgtaggtg cagtgtgtcc gggcccccag ctgcaccggg cactatcctt ggaccctggt 480
ggcagacagc tgaagggtgc ggaccggaac tttcagctga ggcaaaacct ctacctggtg 540
ggcctttggca aggcgtgtgt gggatatggc gctgcagctg aggaactact gggccagcat 600
cttgtgcagg gcgtgatcag cgttcccaag gggatccgtg ctgccatgga gcgtgccggc 660
aagcaggaga tgctgtgaa gccacatagc cgtgtccagg tattcgaggg tgcggaggac 720
aacctcccgg accgcgatgc gctgcgggct gcactggcca tccagcaact ggctgaggga 780
ctcacagctg atgacctgct gctcgtgctg atctcagggt ggggcacacc tgctgcacac 840
agggatgaca ggtaccaatg tcatggacac ccacctctt ttcttgcggc ctcggtgatg 900
gcataggtea cattttggga gttcagagga ggcctacaag ggcaaggtea gatggcagag 960
caagggttgt cctcagggcc tctctaagcc ttagggcccc tctctcctt ggcttggct 1020
gtttgggttaa ctgtcacctt cactcaggg cctctgctct atatctatt ccttcagcc 1080
agactggcag atgggggctt cccctaccc ctgaggatga ggacaaggcc ctcgccagt 1140
tcagcgttcc cgtgcttctc ccttgggcag cctctctctt gagccccca cctgtttct 1200
ttctgtgaag cgagaatgtc tgaaaataaa taggaccatg ccaaaaaaaaa aaaaaaaaaa 1260
aaagg 1265

```

&lt;210&gt; 128

&lt;211&gt; 703

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7509049CB1

&lt;400&gt; 128

```

gttccgaata ccctcgcca cacctggcct tctccatgct cggaaataact tctgcagcg 60
accaacaggc taaagagggg gaaggctctg agggatccag caccggctcc tctccggca 120
accacggtgg gagcgggcga ggaaatggac ataaaccggg gtgtgaaaag ccagggaatg 180
aagcccgcgg gagcggggaa tctgggattc agaactctga gacgtctcct gggatgttta 240
actttgacac tttctggaag aattttaaat ccaagctggg ttcatcaac tgggatgcca 300
taacaagaa cagggtccc ccccccagca cccgagccct cctctacttc agccgactct 360
gggaggattt caaacagaac actcctttcc tcaactggaa agcaattatt gaggggtgcg 420
acgcgtcatc actgcagaaa cgtgcaggca gagccgatca gccgggtgca ggatggcagg 480

```



```

aggtggcagc tgtaacttcc aagaactaca attacaacca gcatgcgtat cccactgcct 540
atgggtgggaa gtactcagtc aagacccttg caaagggggg agtctcacct tcttctcgg 600
cttcccgggt gcaacctggc ctgctgcagt ggggtgaagtt ttggtagggt agtgtcagag 660
tgagccgacc caggccacat cctggcagtg gaggcacagt cac 703

```

```

<210> 129
<211> 1165
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7509086CB1

```

```

<400> 129
ccgacgcgac ccgcgcgcgcg tccgcgcgcgcg ggagttgttg ctgccgcgat gctggtggcg 60
gcggccgcgcg agcgggaacaa ggatcccac tttgcacgtgc tgcggcagta cctggatccg 120
gcccagcgtg gcgtccgcgt cctcgagggtg gcctcgggct ccggccagca cgcagcgcac 180
ttcgcgcgcgcg ccttccccct ggccgagtg cagccgtcgg acgtggacca gcgctgctg 240
gacagcatcg cggccaccac gcaagcccag ggctgacca acgtgaaggc cccgctacac 300
ctggacgtga cgtggggctg ggagcactgg ggcggggatcc tgccacagtc gctggacctg 360
ttgctctgca tcaacatggc ccatgtcagc cccctgcgct gcacggaggg gctcttcaga 420
gcagcaggac acctgctcaa acccagggcc ctgctcatca cctacggggt gagtgggcct 480
gtccacggcg ggctccccgc tcccagtcct agctgctgtc tctgcccctg ctgagagtcc 540
tccctcctgg gctccctgcc ccccgccag ccaccagcac ggctccccct ccacttcag 600
tctgtacccc tccctgggtc atctgccagc ctccctaccc ccttgccac caggctggcc 660
ttggtgact cctcttggg gacagggagg ggggtgccctt tcattctcca aacccagcc 720
tgcaagtctg tcccaggctc tctgggaaat accccgacct tgggtcagtg ggctgtgtcc 780
ccatgcccac agccctatgc catcaatggg aagatctccc ccagagcaa cgtggacttt 840
gacctgatgc tcagatgcag gaaccagaa tgggggcttc gggacacagc cctcctggag 900
gacctgggaa aggcagtg cctgctcctg gagaggatgg tggacatgcc agcaacaac 960
aaatgcctga tcttcggaa aaactaagcc cctccttcac ccccgcacac ctgcatccct 1020
gccggaggct ctgtgaggca cgaaccctgc ctccctaggg cggaccttgt ggacgacagc 1080
cccaccagt ctgtgtctc agccgctggc cgaaggggcc agcctgtctc gaataaagca 1140
tgtcctgctg ccggcaaaaa aaaaa 1165

```

```

<210> 130
<211> 1625
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7506914CB1

```

```

<400> 130
gggagaaggc aggggaggtg gctggtttaa ggggaacttg agggaagtag ggaagactcc 60
tcttgggacc tttggagtag gtgacacatg agcccagccc cagctcacct gccaatccag 120
ctgaggagct cacctgccaa tccagctgag gctgggcaga ggtgggtgag aagagggaaa 180
attgcaggga cctccagttg ggccaggcca gaagctgctg tagctttaac cagacagctc 240
agacctgtct ggaggctgcc agtgacaggt taggtttagg gcagagaaga agcaagacca 300
tggtggggaa gatgtggcct gtgttggtga cactctgtgc agagcatttg ggaaataaaa 360
gtacacatag aagaaatctc agattgcccc acgatctcag aatccacagt cagggtgacc 420
gtcgatgcca tctctgtgga aactccgcag gacgttcttc gggcttcgca gggaaagagt 480
gtcacctgc cctgcaccta ccacacttcc acctccagtc gagagggact tattcaatgg 540
gataagctcc tcctcactca tacggaaagg gtggtcatct ggccgttttc aaacaaaaac 600
tacatccatg gtgagcttta taagaatcgc gtcagcatat ccaacaatgc tgagcagtc 660
gatgcctcca tcaccattga tcagctgacc atggctgaca acggcaccta cgagtgttct 720
gtctcgctga tgtcagacct ggagggcaac accaagtcac gtgtccgcct gttggtctc 780
gtgccaccct ccaaaaccaga atgcggcatc cactccagtc gagaggagaga gaacaacatc 840
cagctgacct gccaatcaaa ggagggtc ccaacccctc agtacagctg gaagaggtac 900
aacatcctga atcaggagca gccctggcc cagccagcct caggtcagcc tgtctccctg 960
aagaatatct ccacagacac atcgggttac tacatctgta cctccagcaa tgaggagggg 1020

```

```

acgcagttct gcaacatcac ggtggccgctc agatctccct ccatgaacgt ggccctgtat 1080
gtgggcatcg cgggtgggctg ggttgccagcc ctccattatca ttggcatcat catctactgc 1140
tgctgctgcc gaggggaagg cgacaacact gaagacaagg aggatgcaag gccgaaccgg 1200
gaagcctatg aggagccacc agagcagcta agagaacttt ccagagagag ggaggaggag 1260
gatgactaca ggcaagaaga gcagaggagc actgggcgtg aatccccgga ccacctcgac 1320
cagtgcacagg ccagcagcag agggcgggcg aggaagggtt aggggttcat tctcccgctt 1380
cctggcctcc cttctccttt ctaagccctg ttctcctgtc cctccatccc agacattgat 1440
ggggacattt cttccccagt gtcagctgtg gggaacatgg ctggcctggt aaggggggtcc 1500
ctgtgctgat cctgctgacc tcaactgtct gtgaagtaac ccctcctggc tgtgacacct 1560
ggtgcggggc tggccctcac tcaagaccag gctgcagcct ccacttccct cgtagtggc 1620
aggag 1625

```

&lt;210&gt; 131

&lt;211&gt; 843

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 5606114CB1

&lt;400&gt; 131

```

gcttccggaa gttgtgagct ttgtttcgga gcgggagctc ttccgagacg cactgggggc 60
cggatgaatc ctgcttatct gtgaaatgca gttaacacat cagctggacc tatttcccgga 120
atgcagggta acccttctgt tatttaaaga tgtaaaaaat gcgggagact tgagaagaaa 180
ggccatggaa ggcaccatcg atggatcact gataaatcct acagtgtttc actcttggtg 240
cccaggctgg agtgcaatgg cgcgatcttg gctcacggca acctctgcct cccgggttca 300
agcaattgtc ctgcctcagc ctctgagtt gctgggatta cagattgttg atccatttca 360
gatacttggt gcagcaaaaca aagcagttca cctctacaaa ctgggaaaaa tgaagacaag 420
aactctatct actgaaatta ttttcaacct ttcccaaat aacaatattt cagaggcttt 480
gaaaaaattt ggtatctcag caaatgcacac ttcaattcta attgtttaca ttgaagaggg 540
agaaaaacaa ataaatcaag aatacctaata atctcaagta gaaggtcatc aggtttctct 600
gaaaaatctt cctgaaataa tgaatattac agaagtcaaa aagatatata aactctcttc 660
acaagaagaa agtattggga cattattgga tgctatcatt tgtagaatgt caacaaaaga 720
tgttttatga aatgtcagaa atattaacaa aaattctcag cattaaagaa aacattgatt 780
ttcctttcct gactataaaa ctaattgtgc attatagaaa agtttaaatac acagaatggt 840
att 843

```

&lt;210&gt; 132

&lt;211&gt; 1589

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7503282CB1

&lt;400&gt; 132

```

cgcttctctc tggccccgagc caggtcagtc cggcagtgca gatggcccat cccgggctct 60
gggctcctag accgggggtgg gccctgccag gagtgggcct gagcctcgca cgtcctctga 120
aactacgcta tgctctgctt ttgcggcgcg atgggcgtgg gagcgggggc tgcccaaaga 180
tgggctgggg ttggaggaag tggcccggta gccgctgtgt gtcctgaggc cgtaaggcgc 240
ctgacccac ccctgaagct ggcgacggag cagggcctgt tttctctaca ctatagtcac 300
gatccgctgg gcccccagcg catctcctgg aagagcccac tcaccctgga cgagctcttc 360
ggtagcctca gaccgtcctt gaagaggatg actgagacat tatggggccac gcgctgtgtg 420
tctgctctcg gggaaactgtc atcattgaca ataagcgcta cctcttcac cagaaactgg 480
gggagggaga ggtacgctgt ggaatgagat agaaaggctg aaggacaaag gcaacttcct 540
gaccgaggat caaatccttt ggctgctgct ggggatctgc agaggccttg aggccattca 600
tgccaaggg ttatgccaca gagacttgaa gcccaccaat atattgcttg gagatgaggg 660
cgagccagtt ttaatggact tgggttccat gaatcaagca tgcattccatg tggagggtct 720
ccgcaggct ctgaccctgc aggactgggc agcccagcgg tgcaccatct cctaccgagc 780
cccagagctc ttctctgtgc agagtcactg tgtcatcgat gagcggactg atgtctggtc 840
cctaggctgc gtgctatatg ccatgatgtt tggggaaggc ccttatgaca tgggtgttcca 900

```

```

aaaggggtgac agtgtggccc ttgctgtgca gaaccaactc agcatccccc aaagccccag 960
gcattcttca gcattgcggc agctcctgaa ctcgatgatg accgtggacc cgcacagcg 1020
tcctcacatt cctctcctcc tcagtcagct ggaggcgctg cagccccag ctcttgcca 1080
acatactacc caaatctgaa aaagcagcat gttgagaaga tggcccttg tgccttgaa 1140
agaggttccc atccctcatt ggaatcacca cccattccat ccaggacttc tcttacactt 1200
gagggtagcg gggtcaggac aatcatctca gtccctgcac ttttcttctg ctttcttccc 1260
tccaagagca aaacctgggc aaggggactt actgagtggg ggtgggtggg ggttgggaaa 1320
agggaaactg gtgggatatg gaacatggct ctgagcagga ctgttgagct cacatagtgt 1380
tctgactcca aatctgggag caggagaatg tgtaaacaag aataaagtgg aagcaggttg 1440
gtgtagatct tagtctcagt gttcctggag tgagagaaga ggtgcaaggg gcccaaagtc 1500
aggttgccat ttcaggcgaa ctgcctggga aggaggaatg gtgcagagtc ggcaatggtg 1560
tgacttcatt ctacaggagg ttggttaga 1589

```

<210> 133  
 <211> 1748  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7503284CB1

```

<400> 133
cgcttctctc tggcccgagc caggtcagtc cggcagtgca gatggcccat cccgggctct 60
gggctcctag accgggggtg gccctgccag gagtgggcct gagcctcgca cgtcctctga 120
aactacgcta tgctctgctt ttgcggcgcg atgggcgtgg gagcgggggc tgcccaaaga 180
tgggctgggg ttggaggaag tggcccggtg gccgctgtgt gtccctgaggc cgtaaggcgc 240
ctgaccccac ccctgaagct ggcgacggag cagggcctgt tttctctaca ctatagtcac 300
gatccgctgg gccccagcg catctcctgg aagagcccac tcaccctgga cgagctcttc 360
ggtagcctca gaccgtcctt gaagaggatg actgagacat tatgggccac gcgctgtgtg 420
tctgctctcg gggaaactgtc atcattgaca ataagcgcta cctcttcac cagaaactgg 480
gggagggcga atcctgtgtc acgagcagca ggaccgggag gaggcccagc gagaagccga 540
catgcatcgc ctcttcaatc accccaacat ccttcgcctc gtggcttact gtctgaggga 600
acgggggtgct aagcatgagg cctggctgct gctaccattc ttcaagagag gtacgctgtg 660
gaatgagata gaaaggctga aggacaaagg caacttcctg accgaggatc aaatcctttg 720
gctgctgctg gggatctgca gaggccttga ggccattcat gccaaagggt atgcccacag 780
agacttgaag cccaccaata tattgcttgg agatgagggg cagccagttt taatggactt 840
gggttccatg aatccaagcat gcatccatgt ggagggctcc cgccaggctc tgaccctgca 900
ggactgggca gccagcggg gcaccatctc ctaccgagcc ccagagctct tctctgtgca 960
gagtcactgt gtcacgatg agcggactga tgtctggtcc ctaggctgcg tgctatatgc 1020
catgatgttt ggggaaggcc cttatgacat ggtgttccaa aagggtgaca gtgtggccct 1080
tgctgtgcag aaccaactca gcatcccaca aagccccagg cattcttcag cattgcggca 1140
gctcctgaac tcgatgatga ccgtggaccc gcatcagcgt cctcacattc ctctcctcct 1200
cagtcagctg gaggcgctgc agccccagc tcctggccaa catactacc aaatctgaaa 1260
aagcagcatg ttgagaagat ggccccttgt gccttggaaa gaggttccca tccctcattg 1320
gaatcaccac ccattccatc caggacttct cttacacttg agggtagcgg ggtcaggaca 1380
atcatctcag tcctgcatct tttcttctgc tttcttccct ccaagagcaa aacctgggca 1440
aggggactta ctgagtgggg gtgggtgggg gttgggaaaa gggaaactgg tgggatatgg 1500
aacatggctc tgagcaggac tgttgagctc ctgactccaa atctgggagc 1560
aggagaatgt gtaaacaaga ataaagtgg agcaggttgg tgtagatctt agtctcagtg 1620
ttcctggagt gagagaagag gtgcaagggg ccaaagtca ggttgccatt tcaggcgaac 1680
tgcctgggaa ggaggaatgg tgcagagtcg gcaatggtgt gacttcattc tacaggaggt 1740
tggttaga 1748

```

<210> 134  
 <211> 621  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7510501CB1

```

<400> 134
gaagattcgg cacaggaaga cgggggattgg ccatcgcttc cgggaagtgt gagctttgtt 60
tcggagcggg agctcttccg agacgcactg ggggcgggat gaatcctgct tatctgtgaa 120
atgcagttaa cacatcagct ggacctattt ccggaatgca ggattgttga tccatttcag 180
atacttgtgg cagcaaaaca agcagttcac ctctacaaac tgggaaaaat gaagacaaga 240
actctatcta ctgaaattat tttcaacctt tccccaaata acaatatttc agaggctttg 300
aaaaaatttg gtatctcagc aaatgacact tcaattctaa ttgtttacat tgaagaggga 360
gaaaaacaaa taaatcaaga atacctaata tctcaagtag aagggtcatca ggtttctctg 420
aaaaatcttc ctgaaataat gaatattaca gaagtcaaaa agatatataa actctcttca 480
caagaagaaa gtattgggac attattggat gctatcattt gtagaatgtc aacaaaagat 540
gttttatgaa atgtcagaaa tattaacaaa aattctcagc attaaagaaa acattgattt 600
tcaaaaaaaa aaaaaaaaaa g                                     621

```

<210> 135

<211> 2220

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7500444CB1

```

<400> 135
cgcgccacgc tctccccgga tcgtgcgggg cctgagcctc tccgccggcg caggctctgc 60
tcgcgccacg tcgctccccg agccatgccc accaccatcg agcgggagtt cgaagagttg 120
gatactcagc gtcgctggga gccgctgtac ttgatgatca cagtcgtgtt aaactgcaaa 180
atgctgagaa tgattatatt aatgccagtt tagttgacat agaagaggca caaaggagtt 240
acatcttaac acagggtcca ctctctaaca catgctgcca tttctggctt atgggtttggc 300
agcagaagac caaagcagtt gtcatgctga accgcattgt ggagaaagaa tcggttaaatt 360
gtgcacagta ctggccaaca gatgaccaag agatgctgtt taaagaaaaca ggattcagtg 420
tgaagctctt gtcagaagat gtgaagtcgt attatacagt acatctacta caattagaaa 480
atatcaatag tggtgaaacc agaacaatat ctacatttca ttatactacc tggccagatt 540
ttggagtccc tgaatcacca gcttcatttc tcaatttctt gtttaaagtg agagaatctg 600
gctccttgaa cctgaccat gggcctgcgg tgatccactg tagtgcaggc attgggcgct 660
ctggcacctt ctctctggta gacacttgct ttgttttgat ggaaaaagga gatgatatta 720
acataaaaaca agtggtactg aacatgagaa aataccgaat ggggtcttatt cagaccccag 780
atcaactgag attctcatat atggctataa tagaaggagc aaaatgtata aaggggagatt 840
ctagtataca gaaacgatgg aaagaacttt ctaagggaaga cttatctcct gcctttgac 900
attcaccaaa caaaaataatg actgaaaaat acaatgggaa cagaataggt ctagaagaag 960
aaaaactgac aggtgaccga tgtacaggac tttctcttaa aatgcaagat acaatggagg 1020
agaacagtga gagtgtctta cggaaacgta ttcgagagga cagaaaaggcc accacagctc 1080
agaagggtgca gcagatgaaa cagaggctaa atgagaatga acgaaaaaga aaaagggtggt 1140
tatattggga acctattctc actaagatgg ggtttatgtc agtcattttg gttggcgctt 1200
ttgttggtcg gacactgttt tttcagcaaa atgccctata aacaattaat tttgccagc 1260
aagcttctgc actagtaact gacagtgtct cattaatcat aggggtttgt ctgcagcaaa 1320
cgctcatat cccaaaaacg gtgcagtaga atagacatca accagataag tgatatttac 1380
agtcacaagc ccaacatctc aggactcttg actgcagggt cctctgaacc ccaaactgta 1440
aatggctgtc taaaataaag acattcatgt ttgttaaaaa ctggtaaaatt ttgcaactgt 1500
attcatacat gtcaaacaca gtatttcacc tgaccaacat tgagatatcc ttatcacag 1560
gatttgtttt tggaggctat ctggatttta actgcactt gatataagca ataaattattg 1620
tggttttatc tacgttattg gaaagaaaat gacatttaaa taatgtgtgt aatgtataat 1680
gtactattga catgggcatc aacactttta ttcttaagca tttcagggtg aatatatttt 1740
ataagtatct atttaatctt ttgtagttaa ctgtactttt taagagctca atttgaaaaa 1800
tctgttacta aaaaaataaa ttgtatgtcg attgaattgt actggatata ttttccattt 1860
ttctaaagag aagtttgata tgagcagtta gaagttggaa taagcaattt ctactatata 1920
ttgcatttct tttatgtttt acagttttcc ccatttttaa aagaaaagca aacaaaagaa 1980
caaaagtttt tcctaaaaat atctttgaag gaaaattctc cttactggga tagtcaggta 2040
aacagttggg caagactttg taaagaaatt gggttctgta aatcccatTA ttgatagtgt 2100
tatttttcat gaaaatttca atgtagtgtg ggtagattat gatttaggaa gcaaaagtaa 2160
gaagcagcat tttatgattc ataatttcag tttactagac tgaagttttg aagtaaacc 2220

```

<210> 136

<211> 1990

<212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7510297CB1

<400> 136  
 gcctacgagt ggcggcgcggc ctgaagggggc acgcgggggga cctgcaaagc tagtgagggg 60  
 cggggcagggc ggcgcgggtgg gggcggggccg agcccggagg ccagatgagc ggacacagcc 120  
 ccacgcgcggg ggccatgcag gtaagtggct cccgacggcc ccacttgaat ttcgatccca 180  
 gaccgggtcc ggcgcctcc ggggcccag cttagcgcg tgctgcagtg gggccgcctg 240  
 acccaaagcg aaaccgaaag ccccgcgag ggtgacctga cgactttccc gggactggaa 300  
 gggggagtc tgcgagagac taggtggcca tgaacggtaa ggcccga aa gaggcgggtgc 360  
 agactgcggc taaggaactc ctcaagtccg tgaaccggag tccctctcct ttccatgctg 420  
 tggctgaatg ccgcaaccgc cttctccagg ttggcttcag tgaactcaag gagactgaga 480  
 aatggaatat taagcccgag agcaagtact tcatgaccag gaactcctcc accatcatag 540  
 cttttgctgt agggggccag tacgttccctg gcaatggctt cagcctcatc gggggccaca 600  
 cggacagccc ctgcctccgg gtgaaacgtc ggtctcgccg cagccagggtg ggcttccagc 660  
 aagtcgggtg ggagacctat ggtggtggga tctggagcac ctggtttgac cgtgacctga 720  
 ctctggctgg acgcgtcatt gtcaagtgcc ctacctcagg tcggctggag cagcagctgg 780  
 tgcacgtgga ggcggccatt cttcgcatcc cacacctggc catccatctg cagcgaaata 840  
 tcaacgagaa ctttgggccc aacacagaga tgcactagt cccattctt gccacagcca 900  
 tccaggagga gctggagaag gggactcctg agccaggggc tctcaatgct gtggatgagc 960  
 ggcaccattc ggtcctcatg tccctgctct gtgcccattc ggggctgagc cccaaggaca 1020  
 tagtggagat ggagctctgc cttgcagaca cccagcctgc ggtcttgggt ggtgcctatg 1080  
 atgagttcat ctttgcctct cggtgggaca atctgcacag ctgcttctgt gccctgcagg 1140  
 ccttgataga ttctgtgca ggccctggct ccctggccac agagcctcac gtgcgcatgg 1200  
 tcacactcta tgacaacgaa gaggtgggggt ctgagagtgc acagggagca cagtcaactgc 1260  
 tgacagagct ggtgctgcgg cggatctcag cctcgtgcca gcacccgaca gccttcgagg 1320  
 aagccatacc caagtccttc atgatcagcg cagacatggc ccatgctgtg catcccaact 1380  
 acctggacaa gcatgaggag aaccaccggc ctttattcca caagggcccc gtgatcaagg 1440  
 tgaacagcaa gcaacgctat gcttcaaacg cgggtgtcaga ggccctgatc cgagaggtgg 1500  
 ccaacaaagt caaggtcccc ctgcaggatc tcatggtccg gaatgacacc ccctgtggaa 1560  
 ccaccattgg acctatcttg gcttctcggc tggggctgcg ggtgctggat ttaggcagcc 1620  
 cccaactggc catgcactct atccgggaga tggcctgcac cacaggagtc ctccagaccc 1680  
 tcacctctt caagggcttc tttgagctgt tcccttctct aagccataat ctcttagtgg 1740  
 attgagccct cttggaaaga cttctctgcc atccctttgc acctgagagg ggaagtctc 1800  
 agctgagctg aagctggatt attaaagtgg attgtcactc agactctccg tgctacgctt 1860  
 atttgagac tagaggagtg ggagttgagc ctggcttgaa cctttggaac cagaaaagtt 1920  
 ggggagcagg tggaggaggc cacactcctg ggagctgatg gttttaaatc tgggttttaa 1980  
 tctctgaaaa 1990

<210> 137  
 <211> 2164  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7640560CB1

<400> 137  
 cctgcggagc cccagtcggg tggagcacgc caggtgagca cgctgcgca gtcgggcccgt 60  
 aaacaagccc accccttccc tttggcccgc cctcgaggg ctggcccgc ggtgggagg 120  
 gcgagaagga agaggcaggg ggtgggggaa gagggagcac caggaaacagc gcctgcgcgg 180  
 tgggcgtgat ccgggcactt agggcaggat gaacgctgct ttccaagatg gcgacggagg 240  
 gaggagggaa ggagatgaac gagattaaga cccaattcac caccggggaa ggtctgtaca 300  
 agctgctgcc gcactcggag tacagccggc ccaaccgggt gcccttcaac tcgcagggat 360  
 ccaacctgt ccgctctcc ttcgtaaacc tcaacgacca gtctggcaac ggcgaccgc 420  
 tctgcttcaa tgtggccgg gagctgtact tctatatcta caagggggtc cgcaaggctg 480  
 ctgacttgag taaaccaata gataaaagga tatacaaagg aacacagcct acttgtcatg 540  
 acttcaacca cctaacagcc acagcagaaa gtgtctctct cctagtgggc ttttccgcag 600

```

gccaaagtcca gcttatagac ccaatcaaaa aagaaactag caaactttttt aatgaggaaa 660
gactaataga caagtacaga gttacctgtg tcaaattgggt tcccgggttcg gaaagccttt 720
tcttagtagc ccactcgagt gggaacatgt acttatataa tgtggagcac acttgtggca 780
ccacagcccc ccactaccag ctctggaagc agggagagag ctttgccgtg cacacttgca 840
agagcaaate caccaggaac cctctcctta agtggacggg gggcgagggg gccctcaacg 900
agtttgcttt ctcccagat ggcaagtctt tagcgtgcgt gagccaggac ggggtttctgc 960
gggtgttcaa ctttgactca gtggagctgc acggtacgat gaaaagctac tttgggggct 1020
tgctgtgtgt gtgctggagc ccggatggca agtacatcgt gacagggtggc gaggacgact 1080
tgggtgacagt ctggtccttt gtagactgcc gagtaatagc cagaggccac gggcacaaagt 1140
cctgggtcag tgtttagtagc tttgaccctt ataccactag tgtagaagaa ggtgacccta 1200
tggagttagg tggcagcgat gaggacttcc aagaccttct tcattttggc agagatcgag 1260
caaatagtag acagtccagg ctctccaaac ggaactctac agacagccgc cccgtaagtg 1320
tcacgtatcg gtttggttcc gtggggccagg acacacagct ctgtttatgg gaccttacag 1380
aagatatcct tttccctcac caaccctct caagagcaag gacacacaca aatgtcatga 1440
atgccacgag tcttctgtgt ggaagcaatg ggaacagtg tacaacaccc ggaactctg 1500
tgccgcctcc tctgccacgg tccaacagcc tccacattc agcagtctca aatgctggca 1560
gcaaaagcag tgtcatggac ggggccattg ctctgggggt cagcaaattt gcaacacttt 1620
cactacatga ccggaaggag aggcaccacg agaaagatca caagcgaaat catagcatgg 1680
gacacatttc tagcaagagc agtgacaaac tgaatctagt taccaaaacc aaaacggacc 1740
ctgtcaaaaac tctgggaacg cccctgtgtc ctctgaattg agatgttccc ttgttagagc 1800
cgctgatatg taaaaagata gcacatgaga ctctgactgt actaatattt cttgaagact 1860
gtatagtcac tgcttgtcag gagggattta tttgcacatg gggaaggcct ggtaaagtgg 1920
taagttttta tccttaatgc tgcaccagat ctagaacttg aataggtagt gacttttttc 1980
tttttcgtgg gaggggtggg gtgtacaatg aatgtgaatg acacttctta ttcttaatgt 2040
aaatctcaat gcacagagc cataattttg gatactgcat gccatgtaat tctgaatcat 2100
ttgataattt accttagagc atttaaaaaa atataatcaa actaattgcc agccaagtca 2160
gtca

```

&lt;210&gt; 138

&lt;211&gt; 1438

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7506087CB1

&lt;400&gt; 138

```

gtggcttgtg gagtggcgac cgttagtgag gcggttgctg agacagacgc tgaggcgggt 60
aggaggagcc cgagccgtaa gggaagccgt gatgagggcc gtgttgacgt ggagagataa 120
agccgagcac tgtataaatg acatcgcat taaagcctgat ggaactcaac tgattttggc 180
tgccggaagc agattactgg tttatgacac ctctgatggc acctacttcc agcccccaa 240
gggacacaaa gacactgtgt actgtgtggc atatgcgaag gatgggttgt ggtctcctga 300
acagaagtct gtctccaaac acaaatcaag cagcaagatc atctgctgca gctggacaaa 360
tgatggtcag tacctggcgc tggggatgtt caatgggatc atcagcatac ggaacaaaaa 420
tgccgaggag aaagtaaaga tccagcggcc ggggggctcc ctctcgcaa tatggtccat 480
ctgctggaac ctttcaagag aggaacgtaa tgacatcctg gctgtggctg actggggaca 540
gaaagtcttc ttctaccagc tgagtggaaa acagattgga aaggatcggg cactgaactt 600
tgacccctgc tgcacagct actttactaa aggcgagtac attttgctgg ggggttcaga 660
caagcaagta tctcttttca ccaaggatgg agtgcggtt gggactgttg gggagcagaa 720
ctctgggtg tggacgtgtc aagcgaaacc ggattccaac tatgtgtgtg tcggctgcca 780
ggacggcacc atttcttct accagcttat tttcagcaca gtccatgggc tttacaagga 840
ccgctatgcc tacagggata gcatgactga cgtcattgtg cagcacctga tcaactgagc 900
gaaagtctcg attaaatgca aagagcttgt caagaagatt gccatctaca gaaatcgatt 960
ggctatccaa ctgccagaga aaatcctcat ctatgagttg tattcagagg acttatcaga 1020
catgcattac cgggtaaaagg agaagattat caagaagttt gagtgcaccc tcttgggtgt 1080
gtgtgccaat cacatcatcc tgtgccagga gaaacggctg cagtgcctgt ccttcagcgg 1140
agtgaaggag cgggagtggt agatggagtc tctcattcgt tacatcaagg tgatcgggtg 1200
ccctcctgga agagaaggcc tcttagtggg gctgaagaag atgtacttgt tagtgtattc 1260
attcatattg attgtaaagg attatttttc actcagtact gatgtccttg gaaatcttac 1320
ctggaaacat gtttgcaaaa aacattattg ggtctttcat cttttttctt ggtattacat 1380
atttgttcaa taaaaattaa acacccaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaat 1438

```